

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3732	98.1	3802	9	D78579	D78579 Homo sapien
2	3653	96.3	3762	6	E12854	E12854 cDNA encodi
3	2875	75.8	4377	6	AX329720	AX329720 Sequence
4	2875	75.8	4977	6	HS012767	U12767 Human mitog
5	2689	70.9	2714	6	AF380383	AF380383 Sequence
6	2689	70.9	2714	6	S81243	S81243 CHN=steroi
7	2500	65.9	4400	6	E10601	E10601 cDNA encodi
8	2500	65.9	4400	6	AX587468	AX587468 Sequence
9	2500	65.9	4400	10	RATNOR1	D38530 Rattus norv
10	2474	65.2	5386	10	BC068150	BC068150 Mus muscu
11	2371	62.5	5115	4	HS0011767	AJ0111767 Sus scro
12	1939	51.1	2546	9	HGNUCREC	X89894 H.sapiens m
13	1485	39.2	1884	6	AX587466	AX587466 Sequence
14	1485	39.2	1884	10	AF191211	AF191211 Mus muscu
15	1435	37.8	14689	9	AL358937	AL358937 Human DNA
16	1216	32.1	1619	6	CQ723907	CQ723907 Sequence
17	1022	28.0	1606	4	AX011768	AJ011768 Sus scro
18	959	25.3	17410	9	AL359710	AL359710 Human DNA
19	944	24.9	4275	10	RNRNALRBF	X86003 R.norvegicu

		/product="neuron derived orphan receptor"	
		/protein_id="BA11419.1"	
		/db_xref="GI:1651191"	
		/translation="MPCVQAYSPGPGSSVAAQTVSYSEYTTIIMNPDYTKLTMDLGS	
		TEITATATSLPSISTFVEGYSSVELKPSVYOMORPLIKVEEGRAPSYHHHHHHH	
		HHHHHQHQHQPSIPPPASSPDEVLPTSMYFKSPSPSTPTTPAPFOAGALWDEAL	
		PSAPGCIAPGLDLPKMAVPTVAGARPLFHFKPSPPHPPAPSPAGGHHLLGYDPTAA	
		AALSLPGLAAAGQAALAEHGYGLPLAKRAAPLAFPLGLTPTSPASSLLGSSPS	
		LPSPTSRSSSEGGTCAVCGDNAACQHYGVRTCEGKGFKRTVOKNAKYVCLANKNC	
		PVDKRRNRQCYCFQKCLSVGMVKEVVRTSLKGRGRPLPSKPSPLQOEPSOPSP	
		SPEICMMNALVALDTSYPRDLDSRYCPTDQAAAGTDAEHVQVQVFNLLTASIDVRS	
		WAERLPFTDLPEKQDTLLIESAFLEFVLRLSIRSNTRAEKDFVFCNGLVLRLOCLR	
		GGEWLDLSKDFSLNLSNLDIQALACLALSULMERHGLKPRKRVBELCNKITVSSL	
		KHQSKQQALESFNESEKVLVALVELRKITLGLQIRIFYLKLELDLVSPSSIIDLFLDTL	
		PF"	
		2613..->3802	
		/gene="NOR-1"	
3' UTR			
ORIGIN			
Query Match		98.1%; Score 3722; DB 9; Length 3802;	
Best Local Similarity		99.4%; Pred. No. 0;	
Matches 3779; Conservative		0; Mismatches 15; Indels 8; Gaps 4;	
QY	1	ATAAATGAGTCCGAGAGAGCGAGCGAAACGGCGAGCGCGGAGAGCGGAGTCTCTGCTT	60
DB	1	ATAAATGAGTCCGAGAGAGCGAGCGAAACGGCGAGCGCGGAGAGCGGAGTCTCTGCTT	60
QY	61	CCGGCCCCACCCCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	120
DB	61	CCGGCCCCACCCCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	120
QY	121	CACCGCTCCCTCACTCGCACACACACACAAAGCGCGCACACAGGCTCCG--CACACAC	178
DB	121	CACCGCTCCCTCACTCGAACACACACACAAAGCGCGCACACAGGCTCCGCGCACACAC	180
QY	179	ACTTGGCTCTCCGGCGGCTGACACCCCTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG	238
DB	181	ACTTGGCTCTCCGGCGGCTGACACCCCTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG	240
QY	239	CGCGGAGCTGGACGCCCTCCGGGCTCACTTTGCAAGCTTGAGGCTGCGGCGAGTGCG	298
DB	241	CGCGGAGCTGGACGCCCTCCGGGCTCACTTTGCAAGCTTGAGGCTGCGGCGAGTGCG	300
QY	299	CGTGAGGTGGAAACAGCGCGGCGATCTCTCCCTCTGGTTCACAGCCCAAGCCAGGACGC	358
DB	301	CGTGAGGTGGAAACAGCGCGGCGATCTCTCCCTCTGGTTCACAGCCCAAGCCAGGACGC	360
QY	359	CGCGGAACCTCTCGGCTGTGCTCTCCCATGATCGGGATCGGATCGGATCGGATCGGAT	418
DB	361	CGCGGAACCTCTCGGCTGTGCTCTCCCATGATCGGGATCGGATCGGATCGGATCGGAT	420
QY	419	CTCACCGCTCCGGGAGCGCTGGGCTTGTAACCGCAGCGCTTCCGGGACAGCAGCTGT	478
DB	421	CTCACCGCTCCGGGAGCGCTGGGCTTGTAACCGCAGCGCTTCCGGGACAGCAGCTGT	480
QY	479	GACTCCCCCAGTGCAGATTTCCGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC	538
DB	481	GACTCCCCCAGTGCAGATTTCCGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC	540
QY	539	CGCCACAGCACTCAAAGCCCACTCGGGAAGAGGGGCGCGGAGCGGCGGCGGCTGAGC	598
DB	541	CGCCACAGCACTCAAAGCCCACTCGGGAAGAGGGGCGCGGAGCGGCGGCGGCTGAGC	600
QY	599	CTGGACCCCTTAGCGGTGCGGCGCAGCACTGCGCGGCTTCTGCTTCTGCGGAGCTCGGCTC	658
DB	601	CTGGACCCCTTAGCGGTGCGGCGCAGCACTGCGCGGCTTCTGCTTCTGCGGAGCTCGGCTC	660
QY	659	CTCTACACTCTCAGCCTCCGCTGGAGAGACCCCGAGCCCACTTCAAGCGCGAAGAT	718
DB	661	CTCTACACTCTCAGCCTCCGCTGGAGAGACCCCGAGCCCACTTCAAGCGCGAAGAT	720
QY	719	ACCTTCCAGATATGCTTCCGCTCCAGCCCAATATAGCCCTTCCCTTCCAGTTTCCAGTT	778

DB	721	ACCCTCCAGATATGCGCTTCAAGCCCAATATAGCCCTTCCCTTCCAGGTTCCAGTT	780
QY	779	ATGCGCGCGAGACATACAGCTCGGAATACACACCGAGATCATGAACCCCGACTACACCA	838
DB	781	ATGCGCGCGAGACATACAGCTCGGAATACACACCGAGATCATGAACCCCGACTACACCA	840
QY	839	AGCTGACCATGAGACTTGGAGACCTGAGATCAGGCTACAGCCACCAAGCTCTGCTGCT	898
DB	841	AGCTGACCATGAGACTTGGAGACCTGAGATCAGGCTACAGCCACCAAGCTCTGCTGCT	900
QY	899	GCATCAGTACCTTCTGAGGAGCTTCTGAGCAACTAGCAACTCAAGCTTCTGCTGCT	958
DB	901	GCATCAGTACCTTCTGAGGAGCTTCTGAGCAACTAGCAACTCAAGCTTCTGCTGCT	960
QY	959	ACCAATGTCAGCGGCTTGTATCAAAATGAGAGGCGGCGGCGGCGGCGGCGGCGGCGG	1018
DB	961	ACCAATGTCAGCGGCTTGTATCAAAATGAGAGGCGGCGGCGGCGGCGGCGGCGGCGG	1020
QY	1019	ATCAC	1078
DB	1021	ATCAC	1080
QY	1079	TTCTCTCAGCTTCCAGCGCGGAGGAGGAGTGTCTGCCAGCAGCTTCCATGTACTTCAAG	1138
DB	1081	TTCTCTCAGCTTCCAGCGCGGAGGAGTGTCTGCCAGCAGCTTCCATGTACTTCAAG	1140
QY	1139	AGTCCCGACCGTCCACCCCGGCTTCCCGCGGCTTCCCGCGGCGGCGGCGGCTTATGG	1198
DB	1141	AGTCCCGACCGTCCACCCCGGCTTCCCGCGGCTTCCCGCGGCGGCGGCGGCTTATGG	1200
QY	1199	ACGAGGACTTGGCTTGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1258
DB	1201	ACGAGGACTTGGCTTGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1260
QY	1259	TGAAGCGGCTTCCACCGTGGCGGCGGCTTCCCGCTTCCCTTCCCTTCCCTTCCCTTCC	1318
DB	1261	TGAAGCGGCTTCCACCGTGGCGGCGGCTTCCCGCTTCCCTTCCCTTCCCTTCCCTTCC	1320
QY	1319	CGCGCATCTCCCGCGGCGGCTTCCCGCGGCGGCTTCCCGCTTCCCTTCCCTTCCCTTCC	1378
DB	1321	CGCGCATCTCCCGCGGCGGCTTCCCGCGGCGGCTTCCCGCTTCCCTTCCCTTCCCTTCC	1380
QY	1379	CGCTTGGCGGCTTACGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGGCT	1438
DB	1381	CGCTTGGCGGCTTACGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGGCT	1440
QY	1439	CGCTTGGAGCGGCTTACGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGG	1498
DB	1441	CGCTTGGAGCGGCTTACGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGG	1500
QY	1499	CGCTTCTCGGCTTACGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCG	1558
DB	1501	CGCTTCTCGGCTTACGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCG	1560
QY	1559	TGCGCTTGGCGGCTTACGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGG	1618
DB	1561	TGCGCTTGGCGGCTTACGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGGCTTGGAGCGG	1620
QY	1619	ACAAGCGGCTTCCAGCACTACGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAG	1678
DB	1621	ACAAGCGGCTTCCAGCACTACGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAG	1680
QY	1679	AGAGAACAGTGCAGAAAAATGCAAAATATGTTTGGCTTGGCAATAAAACTGCCAGTAG	1738
DB	1681	AGAGAACAGTGCAGAAAAATGCAAAATATGTTTGGCTTGGCAATAAAACTGCCAGTAG	1740
QY	1739	ACAAGAGAGTGCAGAACCGATGTCTGATGTCTGATTTTTCAGAAAGTGTCTCAGTGTG	1798
DB	1741	ACAAGAGAGTGCAGAACCGATGTCTGATGTCTGATTTTTCAGAAAGTGTCTCAGTGTG	1800
QY	1799	TGGTAAAAAGAGTGTCTGCTACAGATAGTCTGAAAGGAGGAGAGTGTCTGCTTCCCT	1858
DB	1801	TGGTAAAAAGAGTGTCTGCTACAGATAGTCTGAAAGGAGGAGAGTGTCTGCTTCCCT	1860

QY 1859 AACCAAGAGCCATTACACAGAGACCTTCTCAGCCCTCTCCACCTTCTCCTCCATCT 1918
DB 1861 AACCAAGAGCCATTACACAGAGACCTTCTCAGCCCTCTCCACCTTCTCCTCCATCT 1920
QY 1919 GCATGATGAATGCCCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTA 1978
DB 1921 GCATGATGAATGCTCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTA 1980
QY 1979 CAGATATCTGTCCCACTGACAGGCTGTGACGGCAAGATGTGACATGTGCAACAAT 2038
DB 1981 CCAGATATCTGTCCCACTGACAGGCTGTGACGGCAAGATGTGACATGTGCAACAAT 2040
QY 2039 TCTACAACTCTGACAGGCTCAATTGATGTATCCAGAGCTGGGAGAGAAATCCCG 2098
DB 2041 TCTACAACTCTGACAGGCTCAATTGATGTATCCAGAGCTGGGAGAGAAATCCCG 2100
QY 2099 GATTACTGATCTCCCAAGAGAGATCAGACATTAATTTGATTAATCAGGCTTTTGGAGC 2158
DB 2101 GATTACTGATCTCCCAAGAGAGATCAGACATTAATTTGATTAATCAGGCTTTTGGAGC 2160
QY 2159 TGTTCCTCTCAGACTTTCCATCAGGTCAACACTGCTGAAGATAAGTTTGTGTTCTGCA 2218
DB 2161 TGTTCCTCTCAGACTTTCCATCAGGTCAACACTGCTGAAGATAAGTTTGTGTTCTGCA 2220
QY 2219 ATGGACTTCTCTGCATCGACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2278
DB 2221 ATGGACTTCTCTGCATCGACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
QY 2279 TTAAGACTTTTCTTAAATTTTGCAGAGCTGAACCTTTGATATCCAGGCTTAAGCCTGCC 2338
DB 2281 TTAAGACTTTTCTTAAATTTTGCAGAGCTGAACCTTTGATATCCAGGCTTAAGCCTGCC 2340
QY 2339 TGTGACACTGACATGATCACAAGAAAGACATGGGTTAAAGAACCAAGAGAGTCAAG 2398
DB 2341 TGTGACACTGACATGATCACAAGAAAGACATGGGTTAAAGAACCAAGAGAGTCAAG 2400
QY 2399 AGCTATGCAACAGATCAGAGCAGTTTAAAGACACACAGAGTAAGGACAGGCTCTGG 2458
DB 2401 AGCTATGCAACAGATCAGAGCAGTTTAAAGACACACAGAGTAAGGACAGGCTCTGG 2460
QY 2459 AGCCACCGAGTCCAGGCTCCTGGTCCCTGTAGAACTGAGAGAGATCTGCACCCCTGG 2518
DB 2461 AACCCAGAGTCCAGGCTCCTGGTCCCTGTAGAACTGAGAGAGATCTGCACCCCTGG 2520
QY 2519 GCCTCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGTGTCTCCACCTTCCATCAATTG 2578
DB 2521 GCCTCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGTGTCTCCACCTTCCATCAATTG 2580
QY 2579 ACAGCTCTTCTGGACACCTTACCTTCTAATCAGAGCAGTGGAGCAGTGAAGTGCCT 2638
DB 2581 ACAAGCTCTTCTGGACACCTTACCTTCTAATCAGAGCAGTGGAGCAGTGAAGTGCCT 2640
QY 2639 CCTCTCCTAGCACCTTCTGTACGAGCAAGAGGATAGTTTGGAAACCTATCAATTCC 2698
DB 2641 CCTCTCCTAGCACCTTCTGTACGAGCAAGAGGATAGTTTGGAAACCTATCAATTCC 2700
QY 2699 TGTCTCTCTTAAGAGGAAAGCAGCTCCTGTAGAAAGCAAGACCTTCTTTTTTTCTG 2758
DB 2701 TGTCTCTCTTAAGAGGAAAGCAGCTCCTGTAGAAAGCAAGACCTTCTTTTTTTCTG 2760
QY 2759 GCTCTTTTCTTACAACTTAAGCCAGAAATCTTGACAGATATGTGTGGGTTGTGTT 2818
DB 2761 GCTCTTTTCTTACAACTTAAGCCAGAAATCTTGACAGATATGTGTGGGTTGTGTT 2820
QY 2819 TTATATTTAGGCAATTCGGGATGGGTTGGAGGGGTATAGTTCATGAGGGTTTCTAA 2878
DB 2821 TTATATTTAGGCAATTCGGGATGGGTTGGAGGGGTATAGTTCATGAGGGTTTCTAA 2880
QY 2879 GAAATTCCTAACAAAGCATTCTTTGGACAATGCTATCCAGCAGAGAAAAAGGATAATA 2938
DB 2881 GAAATTCCTAACAAAGCATTCTTTGGACAATGCTATCCAGCAGAGAAAAAGGATAATA 2940

QY 2939 TAACTGTTTTAAACCTTTCTGGGAATCCAAATTATAGTTGCTTTGTATTTTAAACAA 2998
DB 2941 TAACTGTTTTAAACCTTTCTGGGAATCCAAATTATAGTTGCTTTGTATTTTAAACAA 3000
QY 2999 GAACACCAAGGTTGTTGCGCCAGGTTAGGATGTGCTTAAAGATTGGTCCCTTGAATA 3058
DB 3001 GAACACCAAGGTTGTTGCGCCAGGTTAGGATGTGCTTAAAGATTGGTCCCTTGAATA 3060
QY 3059 ATGCTTCTGTATCAAAGGTACGTATGTGTGCAAAACAAGGCAAGAACTTCTTTTAATT 3118
DB 3061 ATGCTTCTGTATCAAAGGTACGTATGTGTGCAAAACAAGGCAAGAACTTCTTTTAATT 3120
QY 3119 TCTCTCTCTCTTTATTTTAACTGTAAGATGAGGATTTACCTACAAATCAGACAT 3178
DB 3121 TCTCTCTCTCTTTATTTTAACTGTAAGATGAGGATTTACCTACAAATCAGACAT 3180
QY 3179 GGCACCAATTAATGCTGTTTGCCTTCCATAAACAAGTGCAATTTTAAAGTGTGCT 3238
DB 3181 GGCACCAATTAATGCTGTTTGCCTTCCATAAACAAGTGCAATTTTAAAGTGTGCT 3240
QY 3239 TACTAAGTCTTGTATTATTAATCTCTCTTTATTTCTATATGGAATAAAGAGGAGCTCA 3298
DB 3241 TACTAAGTCTTGTATTATTAATCTCTCTTTATTTCTATATGGAATAAAGAGGAGCTCA 3300
QY 3299 TGTAGCAATGACACGTTAATATCCCTAGCAGAGCTGTGTTCACTTCCCTGTCGATC 3358
DB 3301 TGTAGCAATGACACGTTAATATCCCTAGCAGAGCTGTGTTCACTTCCCTGTCGATC 3360
QY 3359 CCTTCTGAGTATGGCCCATCCAAAGACTTTTAGGCCATTTCTGATGAAACAGATCCCTG 3418
DB 3361 CCTTCTGAGTATGGCCCATCCAAAGACTTTTAGGCCATTTCTGATGAAACAGATCCCTG 3420
QY 3419 CCTGACTCTCAGCTATCCTGAAAGTGGATCAGATTAATAACTGGATTCATGTAACCTG 3478
DB 3421 CCTGACTCTCAGCTATCCTGAAAGTGGATCAGATTAATAACTGGATTCATGTAACCTG 3480
QY 3479 TTTTGGTCTGTCTTATCAACCCACAGAGTTCCTTAACTTCCCTTCACTTATAGTAAC 3538
DB 3481 TTTTGGTCTGTCTTATCAACCCACAGAGTTCCTTAACTTCCCTTCACTTATAGTAAC 3540
QY 3539 TGACTGATATATTCATTGAGAGCGCATTAAGTCACTGATTTTGTATCCCTAGATAAG 3598
DB 3541 TGACTGATATATTCATTGAGAGCGCATTAAGTCACTGATTTTGTATCCCTAGATAAG 3600
QY 3599 AACATGCAATCAGCAGGAACTGGTCAACAGGTAAGCACCAGGCAATAAGGATTTT 3658
DB 3601 AACATGCAATCAGCAGGAACTGGTCAACAGGTAAGCACCAGGCAATAAGGATTTT 3660
QY 3659 TATAGATATAATTTAATTTTGTTA--TTGGTTAAGGAGA-CAATTTTGGAGGACAGCA 3715
DB 3661 TATAGATATAATTTAATTTTGTGTAATTTGGTTAAGGAGACCAATTTTGGAGGACAGCA 3720
QY 3716 AA---TCTTTTAAAAAATAGTATGAATCTAGAAAAGATTTAAAAAATAGTAT 3772
DB 3721 AATCTTCTTTTAAAAAATAGTATGAATGTGAATCTAGAAAAGATTTAAGAAATAGTAT 3780
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794
DB 3781 GAGTGTGAGTACTAGGAAGGAT 3802

RESULT 2
E12854
LOCUS
DEFINITION
cDNA encoding neuron derived orphan receptor, NOR-1.
E12854
ACCESSION
E12854.1 GI:3251686
VERSION
JP 1997084585-A/1.
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3762)

QY 1659 GGGCTGCAAGGGCTTTTCAAGNAGACAGTGCAGAAAAATGCAAAATATGTTGCTGGC 1718
DB 1621 GGGCTGCAAGGGCTTTTCAAGAGAA CAGTGCAGAAAAATGCAAAATATGTTGCTGGC 1680
QY 1719 AAATAAAAACTGCCAGTAGACAAGAGAGAGTGCAGAACCGATGTCAGTACTGTCGATTTCA 1778
DB 1681 AAATAAAAACTGCCAGTAGACAAGAGAGAGTGCAGAACCGATGTCAGTACTGTCGATTTCA 1740
QY 1779 GAAGTGTCTCAGTGTGGAAATGTAAGAAAGTGTGTCGTAACAGATGATGTCGAAAGGAG 1838
DB 1741 GAAGTGTCTCAGTGTGGAAATGTAAGAAAGTGTGTCGTAACAGATGATGTCGAAAGGAG 1800
QY 1839 GAGAGTGTCTGCTCTCCAAACCAAGAGCCATTACACAGGACCTTCTCAGCCCTC 1898
DB 1801 GAGAGTGTCTGCTCTCCAAACCAAGAGCCATTACACAGGACCTTCTCAGCCCTC 1860
QY 1899 TCCACCTTCTCTCCCAATCTGCATGATGAATGCCCTTGTCCGAGCTTTAAACAGACTCAAC 1958
DB 1861 TCCACCTTCTCTCCCAATCTGCATGATGAATGCTCTTGTCCGAGCTTTAAACAGACTCAAC 1920
QY 1959 ACCAGAGATCTTGATTTTCCAGATATCTGTCCTCACTGACACAGGCTGCTGCAGGACAGA 2018
DB 1921 ACCAGAGATCTTGATTTTCCAGATATCTGTCCTCACTGACACAGGCTGCTGCAGGACAGA 1980
QY 2019 TGCTGAGCATGTGCACAAATTTACAACTCTGACAGCCCTCCATTGATGTATCCAGAAG 2078
DB 1981 TGCTGAGCATGTGCACAAATTTACAACTCTGACAGCCCTCCATTGATGTATCCAGAAG 2040
QY 2079 CTGGGAGAAAGATTTCCGGGATTTACTGATCTCCCAAGAGATGACAGACATTAAT 2138
DB 2041 CTGGGAGAAAGATTTCCGGGATTTACTGATCTCCCAAGAGATGACAGACATTAAT 2100
QY 2139 TGAATCAGCCTTTTGGAGCTGTTGTCTCAGACTTTCCATCAGGTCAAACA CTGCTGA 2198
DB 2101 TGAATCAGCCTTTTGGAGCTGTTGTCTCAGACTTTCCATCAGGTCAAACA CTGCTGA 2160
QY 2199 AGATAGTTTGTCTGCAATGGACTTGTCTGCACTGACACTTCACTGCTGCGATT 2258
DB 2161 AGATAGTTTGTCTGCAATGGACTTGTCTGCACTGACACTTCACTGCTGCGATT 2220
QY 2259 TGGGAGTGGCTGCACTTATTAAGACTTTTCTTAAATTTGACAGCTGAACCTTGA 2318
DB 2221 TGGGAGTGGCTGCACTTATTAAGACTTTTCTTAAATTTGACAGCTGAACCTTGA 2280
QY 2319 TATCCAAGCTTAGCCTGCTGTCAGCACTGAGCATGATCACAGAAAGACATGGGTAAA 2378
DB 2281 TATCCAAGCTTAGCCTGCTGTCAGCACTGAGCATGATCACAGAAAGACATGGGTAAA 2340
QY 2379 AGAACCAAGAGAGTGCAGAGCTATGCAACAAGATCACAGCAGTTTAAAGACACCA 2438
DB 2341 AGAACCAAGAGAGTGCAGAGCTATGCAACAAGATCACAGCAGTTTAAAGACACCA 2400
QY 2439 GAGTAAGGACAGGCTCTGGAGCCCAAGCTCCTGAGTCCCTGAGTCCCTGAGTGAAC 2498
DB 2401 GAGTAAGGACAGGCTCTGGAAACCAAGCTCCTGAGTCCCTGAGTGAAC 2460
QY 2499 GAGGAAGATCTGACACCTGGGCTCCAGGCGATCTTCTACCTGAAGCTGGAAGACTTGG 2558
DB 2461 GAGGAAGATCTGACACCTGGGCTCCAGGCGATCTTCTACCTGAAGCTGGAAGACTTGG 2520
QY 2559 GTCTCCACCTTCATCATTTGA CAAGCTCTTCTGGAACACCTTCTTAATCAGGAGC 2618
DB 2521 GTCTCCACCTTCATCATTTGA CAAGCTCTTCTGGAACACCTTCTTAATCAGGAGC 2580
QY 2619 AGTGGAGCAGTGCCTCTCTCTAGCACTGCTGCTGACGAGCAAGGGATAGS 2678
DB 2581 AGTGGAGCAGTGCCTCTCTCTAGCACTGCTGCTGACGAGCAAGGGATAGG 2640
QY 2679 TTTGAAAACCTATCTATTTCTCTCTTAAAGAGAAAAGCAGCTCTCTGTAGAAGCA 2738
DB 2641 TTTGAAAACCTATCTATTTCTCTCTTAAAGAGAAAAGCAGCTCTCTGTAGAAGCA 2700

QY 2739 AAGACTTTCTTTTCTGCTCTTTCTTACACCTAAAGCCAGAAAACTTGCAGAG 2798
DB 2701 AAGACTTTCTTTTCTGCTCTTTCTTACACCTAAAGCCAGAAAACTTGCAGAG 2760
QY 2799 TATTGTGTTGGGTTGTGTTTATATTTAGGCAATTTGGGGATGGGGTGGAGGGGTAT 2858
DB 2761 TATTGTGTTGGGTTGTGTTTATATTTAGGCAATTTGGGGATGGGGTGGAGGGGTAT 2820
QY 2859 AGTTCATGAGGGTTTCTAAGAAAATGCTAAACAAGCACTTTTGGACAATGCTATCCCAG 2918
DB 2821 AGTTCATGAGGGTTTCTAAGAAAATGCTAAACAAGCACTTTTGGACAATGCTATCCCAG 2880
QY 2919 CAGGAAAAAAGGATAATAAATGCTTTTAAACCTTTCTGCGGAATCCAAATTAAT 2978
DB 2881 CAGGAAAAAAGGATAATAAATGCTTTTAAACCTTTCTGCGGAATCCAAATTAAT 2940
QY 2979 TGTCTTGTAATTTAAAAACAAGCAAGCAAGGGTGTTCGCCAGGGTAGGATGTCTTA 3038
DB 2941 TGTCTTGTAATTTAAAAACAAGCAAGCAAGGGTGTTCGCCAGGGTAGGATGTCTTA 3000
QY 3039 AAGATTGTCCTTGAATAATGCTTCTGTAATCAAGGTA CGTATGTGTGTCAAACAAG 3098
DB 3001 AAGATTGTCCTTGAATAATGCTTCTGTAATCAAGGTA CGTATGTGTGTCAAACAAG 3060
QY 3099 GCAGAAACTTCTTTAATTTCTCTTCTTTTAACTTTTAACTGTAAGATGAGG 3158
DB 3061 GCAGAAACTTCTTTAATTTCTCTTCTTTTAACTTTTAACTGTAAGATGAGG 3120
QY 3159 ATTACCTACAAATCAGACATGCAAAAACAATAATGCTGTTTGTCTTCCATAACAAGTGC 3218
DB 3121 ATTACCTACAAATCAGACATGCAAAAACAATAATGCTGTTTGTCTTCCATAACAAGTGC 3180
QY 3219 AATTTTTTAAAGTGTCTTACTAAGTCTGTGTTAATTAATCTCTCTTTATTCTATATGG 3278
DB 3181 AATTTTTTAAAGTGTCTTACTAAGTCTGTGTTAATTAATCTCTCTTTATTCTATATGG 3240
QY 3279 AATAAAAAAGGAGGAGTCATGTTAGCAAAAGCACTGTTAAATCTCTAGCAGGCTGT 3338
DB 3241 AATAAAAAAGGAGGAGTCATGTTAGCAAAAGCACTGTTAAATCTCTAGCAGGCTGT 3300
QY 3339 GTTACCTTCCCTGTCGATCCCTTCTGAGGTATGGCCCATCCAAGACTTTTAGGCCATTC 3398
DB 3301 GTTACCTTCCCTGTCGATCCCTTCTGAGGTATGGCCCATCCAAGACTTTTAGGCCATTC 3360
QY 3399 TTGATGGAACCAAGATCCCTGCTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTATA 3458
DB 3361 TTGATGGAACCAAGATCCCTGCTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTATA 3420
QY 3459 AACTGGATTACATGTAACCTGTTTGGTGTGTTCTATCAACCCCAAGAGTTCCCTAAA 3518
DB 3421 AACTGGATTACATGTAACCTGTTTGGTGTGTTCTATCAACCCCAAGAGTTCCCTAAA 3480
QY 3519 CTTGCTTCAGTTATAGTAACTGACTGTTATATTCATTCAAGAGCGCCATAAGTCAGTTGA 3578
DB 3481 CTTGCTTCAGTTATAGTAACTGACTGTTATATTCATTCAAGAGCGCCATAAGTCAGTTGA 3540
QY 3579 GTATTGATCCCTAGATAAGAA CAGCAAAATCAGCAGGAACCTGGTCAACAGGGTAAGCA 3638
DB 3541 GTATTGATCCCTAGATAAGAA CAGCAAAATCAGCAGGAACCTGGTCAACAGGGTAAGCA 3600
QY 3639 CCAGGACAAATAGGATTTTATAGATATAATTTAATTTTGTTA - TTGGTTTAAAGAGA 3696
DB 3601 CCAGGACAAATAGGATTTTATAGATATAATTTAATTTTGTGTAATGGGTTTAAAGAGA 3660
QY 3697 -CAATTTTGGAGAGCAAGCAAA - - - TCTTTTTAAAAAATAGTATGTAATGTGAATACTAGA 3752
DB 3661 CCAATTTTGGAGAGCAAGCAAAATCTCTTTTAAAAAATAGTATGTAATGTGAATACTAGA 3720
QY 3753 AAAGATTTTAAAAAATAGTATGAGTGTAGTACTAGGAAGGAT 3794
DB 3721 AAAGATTTTAAAGAAATAGTATGAGTGTAGTACTAGGAAGGAT 3762

```
RESULT 3
AX329720
LOCUS AX329720 4977 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 229 from Patent WO0194629.
ACCESSION AX329720
VERSION AX329720.1 GI:18102698
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrihan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 229 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
1..4977
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 75.8%; Score 2875; DB 6; Length 4977;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;
QY 733 CCTCGCTCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGG--CGCAGA 790
Db |||
94 CTGTGCTGCTCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGTCGTCAGA 153
QY 791 CATACAGCTCGGAATACACCGAGATCATCAACCCGACTACACCAAGCTGACCATGG 850
Db |||
154 CATACAGCTCGGAATACACCGAGATCATCAACCCGACTACACCAAGCTGACCATGG 213
QY 851 ACCTTTGGCAGCACTGAGATCAGGCTACAGCCACCACTCCCTGCCAGCATCAGTACCT 910
Db |||
214 ACCTTTGGCAGCACTGAGATCAGGCTACAGCCACCACTCCCTGCCAGCATCAGTACCT 273
QY 911 TGTGAGGGCTACTCGAGCAATACGAACTCAAGCTTCTCGGTGTACCAATGCGAGC 970
Db |||
274 TTGTGAGGGCTACTCGAGCAATACGAACTCAAGCTTCTCGGTGTACCAATGCGAGC 333
QY 971 GGCCTTGTATCAAGTGGAGGAGGGCGGGCGGCCAGCTTACCATCACCATCACCAACC 1030
Db |||
334 GGCCTTGTATCAAGTGGAGGAGGGCGGGCGGCCAGCTTACCATCACCATCACCAACC 393
QY 1031 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1090
Db |||
394 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 453
QY 1091 CCAGCCCGGAGAGCAGAGTGTCTGCCAGCACTTCAATGTAATTCAGGATGCCACCGT 1150
Db |||
454 CCAGCCCGGAGAGCAGAGTGTCTGCCAGCACTTCAATGTAATTCAGGATGCCACCGT 513
QY 1151 CCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1210
Db |||
514 CCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 573
QY 1211 CCTCGCGCCCGGCTGCATCGCAACCGCGCCGCTGCTGGACCCGCGCATGAAGGGCGTCC 1270
Db |||
574 CCTCGCGCCCGGCTGCATCGCAACCGCGCCGCTGCTGGACCCGCGCATGAAGGGCTCC 633
QY 1271 CCACGGTGGCGCGGCGGCTTCCGCTCTTCCATTCAGGCTTCCGCGCGCATTCGCC 1330
Db |||
634 CCACGGTGGCGCGGCGGCTTCCGCTCTTCCATTCAGGCTTCCGCGCGCATTCGCC 693
QY 1331 CCGCGCCCAACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1390
Db |||
694 CCGCGCCCAACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 753
1391 TCAGCCTCCCGCTGGAGCCGCGAGCCGCGCGGGGAGCCAGCCGCGCGCTGTGAGAGCC 1450
Db |||
754 TCAGCCTCCCGCTGGAGCCGCGAGCCGCGCGGGGAGCCAGCCGCGCGCTGTGAGAGCC 813
QY 1451 ACCCGTACGGGCTGGCGCTGGCCAGAGAGGGGGCGCGCGCTTCCGCGCTCTCGGCG 1510
Db |||
814 ACCCGTACGGGCTGGCGCTGGCCAGAGAGGGGGCGCGCGCTTCCGCGCTCTCGGCG 873
QY 1511 TCACGCCCTCCCGCTACCGCTCCAGCTTGTGGGGGAGAGTCCAGCCTTCCGCTGCGCCG 1570
Db |||
874 TCACGCCCTCCCGCTACCGCTCCAGCTTGTGGGGGAGAGTCCAGCCTTCCGCTGCGCCG 933
QY 1571 CCAGCAGGAGCTCTGCTGTGCGCAGGAGCAGTGTGCCGTGTGCGGGGAGCAACCGCGCT 1630
Db |||
934 CCAGCAGGAGCTCTGCTGTGCGCAGGAGCAGTGTGCCGTGTGCGGGGAGCAACCGCGCT 993
QY 1631 GCCAGCAGTACGGCGTGGAACTTCCGAGGGGTGCAAGGGCTTTTTCAGAGAAACAGTGC 1690
Db |||
994 GCCAGCAGTACGGCGTGGAACTTCCGAGGGGTGCAAGGGCTTTTTCAGAGAAACAGTGC 1053
QY 1691 AGAAAAATGCAAAATATGTTTGCCTGGCAATAAAAACTGCCAGTGTAGACAAGAGACGTC 1750
Db |||
1054 AGAAAAATGCAAAATATGTTTGCCTGGCAATAAAAACTGCCAGTGTAGACAAGAGACGTC 1113
QY 1751 GAAACCGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1810
Db |||
1114 GAAACCGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1173
QY 1811 TTGTCCTGACAGATAGTCTGAAAGGGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1870
Db |||
1174 TTGTCCTGACAGATAGTCTGAAAGGGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1233
QY 1871 CATTAACAAGAGAACCTTCTCAGCCCTCTCAGCCCTCTCAGCCCTCTCAGCCCTCTCAGCCCTCT 1930
Db |||
1234 CATTAACAAGAGAACCTTCTCAGCCCTCTCAGCCCTCTCAGCCCTCTCAGCCCTCTCAGCCCTCT 1293
QY 1931 CCTTTGTCGAGCTTTTAAACAGACTCAACCAAGAGATCTTGAATTTTCCAGATCTGTC 1990
Db |||
1294 CTCTTTGTCGAGCTTTTAAACAGACTCAACCAAGAGATCTTGAATTTTCCAGATCTGTC 1353
QY 1991 CCAGTACAGGCTGCTGAGGAGAGAGTGTGAGCATGTGCAACAAATTTCTACACCTCC 2050
Db |||
1354 CCAGTACAGGCTGCTGAGGAGAGAGTGTGAGCATGTGCAACAAATTTCTACACCTCC 1413
QY 2051 TGACAGCCTTCAATGATGTATCCAGAGCTGGGAGAGAGATTCGGGATTTTCTGATCT 2110
Db |||
1414 TGACAGCCTTCAATGATGTATCCAGAGCTGGGAGAGAGATTCGGGATTTTCTGATCT 1473
QY 2111 TCCCAAGAGAGATCAGACATTTATTTGAATCAAGCCTTTTGGAGCTGTTTGTCTCTCA 2170
Db |||
1474 TCCCAAGAGAGATCAGACATTTATTTGAATCAAGCCTTTTGGAGCTGTTTGTCTCTCA 1533
QY 2171 GACTTTCCATCAGGTCAACACATGCTGAGATTAAGTTTGTGCTTCCATGACATGCTGTC 2230
Db |||
1534 GACTTTCCATCAGGTCAACACATGCTGAGATTAAGTTTGTGCTTCCATGACATGCTGTC 1593
QY 2231 TGCATCGACTTTCAGTGCCTTCTGAGATTTGGGAGTGGCTCGACTCTATTTAAAGACTTTT 2290
Db |||
1594 TGCATCGACTTTCAGTGCCTTCTGAGATTTGGGAGTGGCTCGACTCTATTTAAAGACTTTT 1653
QY 2291 CCTTAAATTTGCAGAGCTGAACTTTGATATCCAAAGCCTTAGCCTGCTGCTGCTGCTGCTGCT 2350
Db |||
1654 CCTTAAATTTGCAGAGCTGAACTTTGATATCCAAAGCCTTAGCCTGCTGCTGCTGCTGCTGCT 1713
QY 2351 GCATCATCACAGAGAGATGGGTTTAAAGAACCAAGAGAGTCCAGAGCTATCCAAACA 2410
Db |||
1714 GCATCATCACAGAGAGATGGGTTTAAAGAACCAAGAGAGTCCAGAGCTATCCAAACA 1773
QY 2411 AGATCAAGCAGATTTTAAAGAACCAAGAGAGTCCAGAGCTCTGAGAGCCACCGAGT 2470
Db |||
1774 AGATCAAGCAGATTTTAAAGAACCAAGAGAGTCCAGAGCTCTGAGAGCCACCGAGT 1833
QY 2471 CCAAGGTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2530
```


[illegible]

```
|||||
2074 AGAGGAAAGCAGCTCCTGTAGAAAGCAAGACTTCTTTTCTGGCTCTTTCCTT 2133
QY
2771 ACAACCTAAAGCCAGAAAACTTGCAGAGTATTGTGTGGGTGTGTTTTATATTAGGC 2830
|||||
2134 ACAACCTAAAGCCAGAAAACTTGCAGAGTATTGTGTGGGTGTGTTTTATATTAGGC 2193
QY
2831 ATTGGGGATGGGTGGGAGGGGTATAGTTTATGTTTCATGNGGGTTTCTAAGAAATTCGTAAC 2890
|||||
2194 ATTGGGGATGGGTGGGAGGGGTATAGTTTATGTTTCATGAGGGTTTCTAAGAAATTCGTAAC 2253
QY
2891 AAAGCACCTTTTGGACAATCTATCCAGCAGGAAAAAAGGATATATAAATGTTTTAA 2950
|||||
2254 AAAGCACCTTTTGGACATCTATCCAGCAGGAAAAAAGGATATATAAATGTTTTAA 2313
QY
2951 AACTCTTCTGGGGAATCCAAATATAGTTTGTGTTTATTTAAACAAAGACGCAAGG 3010
|||||
2314 AACTCTTCTGGGGAATCCAAATATAGTTTGTGTTTATTTAAACAAAGACGCAAGG 2373
QY
3011 GTTGTTCGCAGGTTAGGATGTCTTAAAGATTGGTCCCTTGAAAAATATGCTTCTGTA 3070
|||||
2374 GTTGTTCGCAGGTTAGGATGTCTTAAAGATTGGTCCCTTGAAAAATATGCTTCTGTA 2433
QY
3071 TCAAGGTACGTATGTGGTGCAAAACAGGCAGAACTTCTTTTAAATTTCTTCTTCCTT 3130
|||||
2434 TCAAGGTACGTATGTGGTGCAAAACAGGCAGAACTTCTTTTAAATTTCTTCTTCCTT 2493
QY
3131 TATTTTAAACAAATGTTGAAGATGGAGGATTACCTACAATCAGACATGGCAAAACAATA 3190
|||||
2494 TATTTTAAACAAATGTTGAAGATGGAGGATTACCTAC - AATCAGACATGGCAAAACAAT - 2551
QY
3191 ATGGCTGTTTGTTCATATAACAAAGTGCAATTTTTTAAAGTGCTGTCTTAAGTCTTG 3250
|||||
2552 ATGGCTGTTTGTTC ---ATAACAGCTGCAATTTTTTAAAGTGCTGTCTTAAGTCTTG 2608
QY
3251 TTTATTAACCTCCTTTATTTCTATATGGAATAAAGAGGAGCGATGTTAGCAATG 3310
|||||
2609 T ---TATACTCTCTTATCTATATGCGGAATAAAGAGGAGCGATGTTAGCAATG 2665
QY
3311 ACACGTTAATATCCCTAGCAGAGGCTGTGTTCACCTTCCTCTCGATCCCTTCTCAGGTA 3370
|||||
2666 ACACGTTAATATCCCTAGCAGAGGCTGTGTTCACCTTCCTCTCGATCCCTTCTCAGGTA 2725
QY
3371 TGGCCCATCCAAGCTTTTAGGCAATCTTGTATGGAACAGATCCCTGCGCTGACTGTCC 3430
|||||
2726 TGGCCCATCCAAGCTTTTAGGCAATCTTGTATGGAACAGATCCCTGCGCTGACTGTCC 2785
QY
3431 AGCTATCCTGAAAGTGGATCAGATTATAAAGTGGATTACATGTTTGGTTGTGT 3490
|||||
2786 AGCTATCCTGAAAGTGGATCAGATTATAAAGTGGATTACATGTTTGGTTGTGT 2845
QY
3491 TCTATCAACCCCAACAGAGTTCCTTAACTTGCTTCAGTTATAGTAACCTGACTGGTATAT 3550
|||||
2846 TCTATCAACCCCAACAGAGTTCCTTAACTTGCTTCAGTTATAGTAACCTGACTGGTATAT - 2904
QY
3551 TCATTCAAGAGCCCAATAGTCAAGTGTAGTATTTGATCCCTAGATAAGAACATGCAATC 3610
|||||
2905 TCATTCAAGAGCCCAATAGTCAAGTGTAGTATTT - GATCCTAGATAAGAACATGCAATC --- 2959
QY
3611 AGCAGGAACCTGTCATACAGGTTAGCAACAGGACAATTAAGATTTTTATAGATAATAT 3670
|||||
2960 TCAGAGGACTGTCATACAGGTTAGCAACAGGACAATTAAGATTTTTATAGATAATAT 3019
QY
3671 TTAATTTTGTATTGTTTAAAGGACAATTTTGGAGACAAGCAAA ---TCTTTTAAA 3727
|||||
3020 TTAATTTTGTATTGTTTAAAGGACAATTTTGGAGACAAGCAAAATCTCTCTTTTAAA 3079
QY
3728 AAATAGTATGAATGTGAATCTAGAAAAGATTTTAAAAAATAGTATGAGTGTGAGTACTAG 3787
|||||
3080 AAATAGTATGAATGTGAATCTAGAAAAGATTTTAAAAAATAGTATGAGTGTGAGTACTAG 3139
3788 GAAGGAT 3794
|||||
```

```
Db 3140 GAAGGAT 3146
RESULT 5
AR380383
LOCUS AR380383 2714 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 928 from patent US 6607879.
ACCESSION AR380383
VERSION AR380383.1 GI:40088017
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2714)
AUTHORS Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.
TITLE Compositions for the detection of blood cell and immunological response gene expression
JOURNAL Patent: US 6607879-A 928 19-AUG-2003;
FEATURES
LOCATION/Qualifiers
source 1..2714
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 70.9%; Score 2689.6; DB 6; Length 2714;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2692; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 132 TCACCTCGCACACACAGACACAAGCGCGACACAGGCTCCGCACACACACTTCGCTCTCCC 191
Db 19 TCACACAACACACAGACACAAGCGCGACACAGGCTCCGCACACACACTTCGCTCTCCC 78
QY 192 GCGCGCTCACACCCCTCTTGGCCCTGAGCCCTTGCCTGAGCGGTGACGCGGCGCGAGTGGGA 251
Db 79 GCGCGCTCACACCCCTCTTGGCCCTGAGCGGTGACGCGGCGCGAGTGGGA 138
QY 252 GCGCCCTCCCGGGCTCCTTTGCAAGCTGACGCTGCGGCGAGTGGCGGTGGAGTGGGA 311
Db 139 GCGCCCTCCCGGGCTCCTTTGCAAGCTGACGCTGCGGCGAGTGGCGGTGGAGTGGGA 198
QY 312 ACAGCGGGGCGATCTCCCGCTGGTACAGCGCAAGCCAGACGCGCGGAAACCTCTC 371
Db 199 ACAGCGGGGCGATCTCCCGCTGGTACAGCGCAAGCCAGACGCGCGGAAACCTCTC 258
QY 372 GCGTGTCTCTCCATGAGTCGGGATCGCAGATCCCCCAGCGCTCAGCGCTCGG 431
Db 259 GCGTGTCTCTCCATGAGTCGGGATCGCAGATCCCCCAGCGCTCAGCGCTCGG 318
QY 432 GAGCGCGTGGGCTTGTACACCGAGCCCTTCGGGACAGCAGCTGTGACTCCCCCCCAG 491
Db 319 GAGCGCGTGGGCTTGTACACCGAGCCCTTCGGGACAGCAGCTGTGACTCCCCCCCAG 378
QY 492 TGCAGATTCGGGACAGCTCTCTAGAACTCCGCTCTAAAGACGGAACCGCCACAGCACTC 551
Db 379 TGCAGATTCGGGACAGCTCTCTAGAACTCCGCTCTAAAGACGGAACCGCCACAGCACTC 438
QY 552 AAAGCCCACTCGGGAAGGCGAGCCCGCAAGCCCGGCCCTGAGCCTTGGACCTTACG 611
Db 439 AAAGCCCACTCGGGAAGGCGAGCCCGCAAGCCCGGCCCTGAGCCTTGGACCTTACG 498
QY 612 GTGCGGGGACAGCTGCGCGGCTTTCGCTCGCGGACGTCGCTCTCTACACTCTC 671
Db 499 GTGCGGGGACAGCTGCGCGGCTTTCGCTCGCGGACGTCGCTCTCTACACTCTC 558
QY 672 AGCCTCGCTGAGAGAGCCCCCAGCCCCACCACTTACGCGGCAAGATACCTCCACATAT 731
Db 559 AGCCTCGCTGAGAGAGCCCCCAGCCCCACCACTTACGCGGCAAGATACCTCCACATAT 618
QY 732 GCGCTCGCTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATCGCGCGAGAC 791
Db 619 GCGCTCGCTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATCGCGCGAGAC 678
QY 792 ATACAGCTCGGAATACACACGAGAGATCATGAACCCCGACTACACCAAGCTGACCATGGA 851
```


Db 679 ATACAGCTCGGAATACACACGGAGATCATGAACCCGACTACACCAAGCTGACCATGGA 738
QY 852 CCTTGGCAGCACTGAGATCAGCGCTACAGCCACACAGCTCCCTGCCAGCATCAGTACCTT 911
Db 739 CCTTGGCAGCACTGAGATCAGCGCTACAGCCACACAGCTCCCTGCCAGCATCAGTACCTT 798
QY 912 CGTGAGGGCTACTCGAGCAACTACGAACTCAAGCTTCTCGCTGTATCCAAATCAGCG 971
Db 799 CGTGAGGGCTACTCGAGCAACTACGAACTCAAGCTTCTCGCTGTATCCAAATCAGCG 858
QY 972 GCCCTTGATCAAGTGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1031
Db 859 GCCCTTGATCAAGTGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 918
QY 1032 CCACACACACACACACACACACACACAGCAGCAGCATCAGCAGCAGCATCCATTCTCTCCAGCCTC 1091
Db 919 CCACACACACACACACACACACACACAGCAGCAGCATCAGCAGCAGCATCCATTCTCTCCAGCCTC 978
QY 1092 CAGCCCGGAGGACGAGGTGCTGCCAGCAGCAGCTCCATGTACTTCAAGCAGTCCCGCCCGCTC 1151
Db 979 CAGCCCGGAGGACGAGGTGCTGCCAGCAGCAGCTCCATGTACTTCAAGCAGTCCCGCCCGCTC 1038
QY 1152 CACCCCCACACAGCCCGCTTCCCGCCGAGCAGCGGGGGGGTATGGGACGAGGCACTGCC 1211
Db 1039 CACCCCCACACAGCCCGCTTCCCGCCGAGCAGCGGGGGGGTATGGGACGAGGCACTGCC 1098
QY 1212 CTGCGCGCCGCTGCATCGACACCGCGCGCTGTGGACCGCGCGATGAAGCGGGTCC 1271
Db 1099 CTGCGCGCCGCTGCATCGACACCGCGCGCTGTGGACCGCGCGATGAAGCGGGTCC 1158
QY 1272 CACGCTGCGCGCGCGCTTCCCGCTCTTCCACTTCAAGCCTCGCGCGCGCATCCCGC 1331
Db 1159 CACGCTGCGCGCGCGCTTCCCGCTCTTCCACTTCAAGCCTCGCGCGCGCATCCCGC 1218
QY 1332 CGCGCCACAGCCCGCGCGCGCGCACACCTCGGCTACGACCGAGCGGGCGCTGCCGCT 1391
Db 1219 CGCGCCACAGCCCGCGCGCGCGCACACCTCGGCTACGACCGAGCGGGCGCTGCCGCT 1278
QY 1392 CAGCCTGCGCTGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCTTGGAGGCA 1451
Db 1279 CAGCCTGCGCTGGAGCGCGAGCGCGCGCGCGCGCGCGCGCTTGGAGGCA 1338
QY 1452 CCGCTACGCGCTGCGCTGGCGCAAGAGGGCGCGCGCTTCCGCTTCCGCTTCCGCT 1511
Db 1339 CCGCTACGCGCTGCGCTGGCGCAAGAGGGCGCGCGCTTCCGCTTCCGCTTCCGCT 1398
QY 1512 CAGCGCTCCCTACCGGCTCAGCGCTGCTGGGGGAGAGTCCAGCCTGCGCTGCCGCGC 1571
Db 1399 CAGCGCTCCCTACCGGCTCAGCGCTGCTGGGGGAGAGTCCAGCCTGCGCTGCCGCGC 1458
QY 1572 CAGCAGGAGCTCGTCTGCGGAGGCGACGCTGCTGCGGTCGCGGGACACGCGCGCTG 1631
Db 1459 CAGCAGGAGCTCGTCTGCGGAGGCGACGCTGCTGCGGTCGCGGGACACGCGCGCTG 1518
QY 1632 CAGCAGCTACGCGCTGCGAACTCGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGCA 1691
Db 1519 CAGCAGCTACGCGCTGCGAACTCGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGCA 1578
QY 1692 GAAATATGCAATATGTTGCTCGGCAATATAAACTGCCAGTACAGAGAGCGTCCG 1751
Db 1579 GAAATATGCAATATGTTGCTCGGCAATATAAACTGCCAGTACAGAGAGCGTCCG 1638
QY 1752 AAACCGATGCTAGTACTCTGCAATTTTCAAGAGTGTCTCAGTGTGGAATGTTAAAGAGT 1811
Db 1639 AAACCGATGCTAGTACTCTGCAATTTTCAAGAGTGTCTCAGTGTGGAATGTTAAAGAGT 1698
QY 1812 TGTCCGTACAGATGCTGAAAGGAGGAGGAGTCTCTGCTTCCAAACCAAGAGCGCC 1871
Db 1699 TGTCCGTACAGATGCTGAAAGGAGGAGGAGTCTCTGCTTCCAAACCAAGAGCGCC 1758
QY 1872 ATTACACAGGAGCTTCTCAGCCCTTCCACCTTCTCTCCATCTGCATGATGATGC 1931

Db 1759 ATTACACAGGAACTTCTCAGCCCTCTCCACCTTCTCTCTCAATCTGATGATGATGC 1818
QY 1932 CCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTTATTCAGATCTGTGCC 1991
Db 1819 CCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTTATTCAGATCTGTGCC 1878
QY 1992 CACTGACAGAGCTGTGCGAGGACAGATGCTGAGATGTGGAACAATTTCAACCTCTCT 2051
Db 1879 CACTGACAGAGCTGTGCGAGGACAGATGCTGAGATGTGGAACAATTTCAACCTCTCT 1938
QY 2052 GACAGCTTCCATTTGATGTATCCAGAGCTGGGACAGAAAGATTCGGGATTTACTGATCT 2111
Db 1939 GACAGCTTCCATTTGATGTATCCAGAGCTGGGACAGAAAGATTCGGGATTTACTGATCT 1998
QY 2112 CCCAAAGAGATCAGACATTAATTTGAATCAGCCTTTTGGAGCTTTTGTCTCTCAG 2171
Db 1999 CCCAAAGAGATCAGACATTAATTTGAATCAGCCTTTTGGAGCTTTTGTCTCTCAG 2058
QY 2172 ACTTTCCATCAGGTCAAAACACCTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTGTCTCT 2231
Db 2059 ACTTTCCATCAGGTCAAAACACCTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTGTCTCT 2118
QY 2232 GCATCGACTTCAGTGCCTTCTGGAATTTGGGAGTGGCTCGACTCTATTTAAGAGCTTTTC 2291
Db 2119 GCATCGACTTCAGTGCCTTCTGGAATTTGGGAGTGGCTCGACTCTATTTAAGAGCTTTTC 2178
QY 2292 CTTAAATTTGAGAGCCTGAACCTTGATATCAAGCCTTTAGCCTGCTCTGCTCAGACTGAG 2351
Db 2179 CTTAAATTTGAGAGCCTGAACCTTGATATCAAGCCTTTAGCCTGCTCTGCTCAGACTGAG 2238
QY 2352 CATGATCAAGAGACATGGGTTAAAGAACAAAGAGAGTGAAGAGCTATGCAACAA 2411
Db 2239 CATGATCAAGAGACATGGGTTAAAGAACAAAGAGAGTGAAGAGCTATGCAACAA 2298
QY 2412 GATCAACAGCAGTTTAAAGAACCCAGAGTAAGGAGCAGGCTCTGGAGCCACCGAGTC 2471
Db 2299 GATCAACAGCAGTTTAAAGAACCCAGAGTAAGGAGCAGGCTCTGGAGCCACCGAGTC 2358
QY 2472 CAAGCTCTGGGTGCTGCTGCTAGAACTCAGGAAGATCTGCACTCTGGGCTCTCAGCGCAT 2531
Db 2359 CAAGCTCTGGGTGCTGCTGCTAGAACTCAGGAAGATCTGCACTCTGGGCTCTCAGCGCAT 2418
QY 2532 CTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATATGACAAGCTCTTCTCT 2591
Db 2419 CTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATATGACAAGCTCTTCTCT 2478
QY 2592 SGACACCTTACCTTTTAAATCAGGAGCAGTGGAGCAGTGGCTCTCTCTCTCCTAGCAC 2651
Db 2479 GGACACCTTACCTTTTAAATCAGGAGCAGTGGAGCAGTGGCTCTCTCTCTCCTAGCAC 2538
QY 2652 CTGCTTGTCTAGCAGCAAGGATAGGTTTGGAAACCTATCATATTTCTCTCTCTCTTAA 2711
Db 2539 CTGCTTGTCTAGCAGCAAGGATAGGTTTGGAAACCTATCATATTTCTCTCTCTCTTAA 2598
QY 2712 GAGGAAAGCAGCTCTCTGTAAGAACAAAGACTTCTTTTCTTCTGGCTCTTTCTCTTA 2771
Db 2599 GAGGAAAGCAGCTCTCTGTAAGAACAAAGACTTCTTTTCTTCTGGCTCTTTCTCTTA 2658
QY 2772 CAACCTTAAGCCAGAAAACCTTGACAGTATTTGTTGGGTTGTGTTTATATTTA 2827
Db 2659 CAACCTTAAGCCAGAAAACCTTGACAGTATTTGTTGGGTTGTGTTTATATTTA 2714

RESULT 6

S81243
LOCUS S81243 2714 bp mRNA linear PRI 09-MAY-1996
DEFINITION CHN=steroid/thyroid orphan receptor homolog gene [human, fetal brain, mRNA partial, 2714 nt].
ACCESSION S81243
VERSION S81243.1 GI:1311504
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2714) Clark,J., Benjamin,H., Gill,S., Sidhar,S., Goodwin,G., Crew,J., Gusterson,B.A., Shipley,J. and Cooper,C.S. Fusion of the EWS gene to CHN, a member of the steroid/thyroid receptor gene superfamily, in a human myxoid chondrosarcoma Oncogene 12 (2), 229-235 (1996) 96152889 MEDLINE 8570200 PUBMED GenBank staff at the National Library of Medicine created this entry [NCBI Gibbs 176289] from the original journal article.

FEATURES source 1..2714 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" 1..2714 /gene="CHN" /note="steroid/thyroid orphan receptor homolog gene; rat NOR1 homolog" 443..2497 /gene="CHN" /note="steroid/thyroid orphan receptor homolog gene; No start codon was found. N-terminal of protein truncated by chromosomal breakpoint 2 nucleotides 5' of first (Pro) codon" start=1 /codon_start=1 /protein_id="AAB36006.1" /db_xref="GI:1311505" /translation="PTAEGSPASPEPGPLAVPGSTAGSPRTSAPPTLSASAGE TSPITQIARYPDMPQVQAOYSPRPGSSYAAOTYSSEYTTIEINNPDYKLTMDLGS THETATATLSISFTVEGYSSNELAPSCVYQMRPLIKVEGRASRHHHHHHH HHHHHQHQQSIPPASSDEVLPSFTMYKQSPSTPTTPAPPQAGALWDEAL PSAPCIAPGLDPPMKAVPTVAGARFPLFKPSPHPAPSPAGGHHLYDPTAA AALSLPLGAAAAGQAALSHYPGLPLAKAAAPLAPPLGLTSPSTASSLLGESPS LPSPSRSSSGEGTCAVCNDNAACHQVVRTCEGKGFPRKTVOKNAKYCLANKNC PVDKRRNRCOVCRQKCLSVGMVKEVVRTSLKGRGRPLPSKPSPLQOEPSRPP SPPICMNALVRLIDSTPRDLDSRYCPTDQAAAGTDAEHVQQYNNLTASIDVRS WAEKIPGFTDLPEQDOLLIESAFLEFLVRLISRTAEKDFVFCNGVLVHLRQLR FGFEWDSITKPSLMSQLNLIDIALCLMSLMTERRHLKPRVBEELCNKITSSL KDHQSGQALEPTESKVLGALVELRKICTLGLQIRIFYLKLELDVSPPSIIDKLFDLT LP"

ORIGIN

Query Match 70.9%; Score 2689.6; DB 9; Length 2714; Best Local Similarity 99.9%; Pred. No. 0; Matches 2692; Conservative 0; Mismatches 4; Indels 0; Gaps 0; QY 132 TCACGCGACACACAGCGGCGCACAGGCTCCGACACACACTTCGCTCTCC 191 Db 19 TCACACAAACACACAGCGCGCACAGGCTCCGACACACACTTCGCTCTCC 78 QY 192 GCGCGCTCACACCCCTCTTGCCTCAGCCCTTGCCGGTGACGGCGGCGCGAGCTGGA 251 Db 79 GCGCGCTCACACCCCTCTTGCCTCAGCCCTTGCCGGTGACGGCGGCGAGCTGGA 138 QY 252 CGCCCTCCCGGGCTCACTTTGCAAGCTGACGGTGCGCGGAGTGGCGTGGAGTGGGA 311 Db 139 CGCCCTCCCGGGCTCACTTTGCAAGCTGACGGTGCGCGGAGTGGCGTGGAGTGGGA 198 QY 312 ACAGCGGGGCGATCTCCCTCTGTCTACAGCCCAAGCCAGGACGGCGCGGAACCTCTC 371 Db 199 ACAGCGGGGCGATCTCTCCCTCTGTCTACAGCCCAAGCCAGGACGGCGCGGAACCTCTC 258 QY 372 GCGTGTCTCTCCCATGAGTCGGGATCGGAGATCCCCCAGCGGCTCACCGGCTCG 431 Db 259 GCGTGTCTCTCCCATGAGTCGGGATCGGAGATCCCCCAGCGGCTCACCGGCTCG 318 QY 432 GGAGCGGCTGGGCTTTGTACACCGGAGCCCTTCGGGACAGAGCTGTGTACTCCCGCCAG 491 Db 319 GGAGCGGCTGGGCTTTGTACACCGGAGCCCTTCGGGACAGAGCTGTGTACTCCCGCCAG 378

QY 492 TGCAGATTTTCGGGACAGCTCTCTAGAAACTCGCTCTTAAAGACGGAACCGCCACAGCACTC 551 Db 379 TGCAGATTTTCGGGACAGCTCTCTAGAAACTCGCTCTTAAAGACGGAACCGCCACAGCACTC 438 QY 552 AAAGCCCACTGCGGAAGAGGGGAGCGCCCGCAAGCCCGGCGCTTGAGCCCTTGAGCCCTT 611 Db 439 AAAGCCCACTGCGGAAGAGGGGAGCGCCCGCAAGCCCGGCGCTTGAGCCCTTGAGCCCTT 498 QY 612 GGTGCGGGGACAGCTGCGCGGCGCTTTCGCTCGCGGAGCGTTCGGCTCTCTACACTCTC 671 Db 499 GGTGCGGGGACAGCTGCGCGGCGCTTTCGCTCGCGGAGCGTTCGGCTCTCTACACTCTC 558 QY 672 AGCTTCGCTGAG 731 Db 559 AGCTTCGCTGAG 618 QY 732 GGCCTGCGTCAAGGCCAATATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGCGAGAG 791 Db 619 GGCCTGCGTCAAGGCCAATATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGCGAGAG 678 QY 792 ATACAGCTCGGAATACACACAGGAGATCATGAACCCCGGAGCTACACCAAGCTGACCATGA 851 Db 679 ATACAGCTCGGAATACACACAGGAGATCATGAACCCCGGAGCTACACCAAGCTGACCATGA 738 QY 852 CCTTGGCAGCACTGAGATCAGCGCTACAGCCACCAAGCTCCCTGCGCCAGCATCAGTACCTT 911 Db 739 CTTTGGCAGCACTGAGATCAGCGCTACAGCCACCAAGCTCCCTGCGCCAGCATCAGTACCTT 798 QY 912 CGTGAGGGGCTACTCGAGCAACTAGCAACTCAAGCTTCTCTGCTGTGTACCAATGAGCGG 971 Db 799 CGTGAGGGGCTACTCGAGCAACTAGCAACTCAAGCTTCTCTGCTGTGTACCAATGAGCGG 858 QY 972 GGCCTTGATCAAAGTGGAGAGGGGCGGCGCCAGCTACCATCACCATCAACACACCA 1031 Db 859 GGCCTTGATCAAAGTGGAGAGGGGCGGCGCCAGCTACCATCACCATCAACACACCA 918 QY 1032 CCACACACACACACACACACACACACACAGCAGATCAGCAGCCATCCTTCTCCAGCCTC 1091 Db 919 CCACACACACACACACACACACACACAGCAGATCAGCAGCCATCCTTCTCCAGCCTC 978 QY 1092 CAGCCCGGAGGACGAGGTGCTGCCCGACACCTCCATGTACTTCAAGCAGTCCCGCCAGCTC 1151 Db 979 CAGCCCGGAGGACGAGGTGCTGCCCGACACCTCCATGTACTTCAAGCAGTCCCGCCAGCTC 1038 QY 1152 CACCCCGACACCGCGGCTTCCCGCGAGAGGGGCGGCGCTTATGGGAGAGGACATGCC 1211 Db 1039 CACCCCGACACCGCGGCTTCCCGCGAGAGGGGCGGCGCTTATGGGAGAGGACATGCC 1098 QY 1212 CTCGCGGCGCGGCTGCATCGCACCCGCGCTGTGGACCCGCGATGAAGCGGCTCC 1271 Db 1099 CTCGCGGCGCGGCTGCATCGCACCCGCGCTGTGGACCCGCGATGAAGCGGCTCC 1158 QY 1272 CACGCTGCGCGGCGGCTTCCCGCTCTTTCACCTTCAAGCCTTCCCGCGGATCCCC 1331 Db 1159 CACGCTGCGCGGCGGCTTCCCGCTCTTTCACCTTCAAGCCTTCCCGCGGATCCCC 1218 QY 1332 CGCGCCAGCCCGCGCGGCGGCGCACCACTCGGCTACGACCGAGCGCGCTGCGCGCT 1391 Db 1219 CGCGCCAGCCCGCGCGGCGGCGCACCACTCGGCTACGACCGAGCGCGCTGCGCGCT 1278 QY 1392 CAGCCTGCGCTGCGGAGCGGAGCGCGCGGAGCGGAGCGGCGGCGCTTGGAGGCGCA 1451 Db 1279 CAGCCTGCGCTGCGGAGCGGAGCGCGCGGAGCGGAGCGGCGGCGCTTGGAGGCGCA 1338 QY 1452 CCGCTACCGGCTGCGGCTGGCGCAAGAGGCGGCGCGCTTCCCGCTCTCTCGGCTC 1511 Db 1339 CCGCTACCGGCTGCGGCTGGCGCAAGAGGCGGCGCGCTTCCCGCTCTCTCGGCTC 1398 QY 1512 CAGCCCTCTCCCTACCGGCTCAGGCTGCTGGGCGAGAGTCCAGCTGCGCTGCGCGCC 1571 Db 1399 CAGCCCTCTCCCTACCGGCTCAGGCTGCTGGGCGAGAGTCCAGCTGCGCTGCGCGCC 1458 QY 1572 CAGCAGGAGCTCGTCTGCGGAGGGGACAGTGTGCTGCGGGGACAAAGCGCGCTG 1631

Db	1459	CAGCAGGAGCTCGTGTGTCGCGAGGCGACGCTGTCCGCTGTGCGGGGACAACGCCCCCTG	1518
Qy	1632	CCAGCACTTCGGCGTGCAGAACTTGCAGAGGCTGCAAGGGCTTTTTCAGAGAGAAACAGTGCA	1691
Db	1519	CCAGCACTTCGGCGTGCAGAACTTGCAGAGGCTGCAAGGGCTTTTTCAGAGAGAAACAGTGCA	1578
Qy	1692	GAAAAATGCAAAATATCTTTTCCCTGGCAAAATAAAAACTGCCAGTAGTACACAAGAGACGCTCG	1751
Db	1579	GAAAAATGCAAAATATCTTTTCCCTGGCAAAATAAAAACTGCCAGTAGTACACAAGAGACGCTCG	1638
Qy	1752	AAACCGATGTCAGTACTGTGTCGATTTTCAGAAAGTGTCTCAGTGTGTGGAATGTTAAAGAAAGT	1811
Db	1639	AAACCGATGTCAGTACTGTGTCGATTTTCAGAAAGTGTCTCAGTGTGTGGAATGTTAAAGAAAGT	1698
Qy	1812	TGTCGGTACAGATAGTCTGAAGAGGAGGAGAGTGTCTGCTGCCCTTCCAAACCAAGAGAGCCC	1871
Db	1699	TGTCGGTACAGATAGTCTGAAGAGGAGGAGAGTGTCTGCTGCCCTTCCAAACCAAGAGAGCCC	1758
Qy	1872	ATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCTCCATCTGCAATGATGAATGC	1931
Db	1759	ATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCTCCATCTGCAATGATGAATGC	1818
Qy	1932	CCTTGTCCGAGCTTTTAAACAGACTCAACCAACAGAGATCTTGTATTTTCAGATACTGTCTC	1991
Db	1819	CCTTGTCCGAGCTTTTAAACAGACTCAACCAACAGAGATCTTGTATTTTCAGATACTGTCTC	1878
Qy	1992	CACGTACACAGGCTGCTGCAGGCAACAGATCTGAGCATGTGCAACATTTCTACAACTCCT	2051
Db	1879	CACGTACACAGGCTGCTGCAGGCAACAGATCTGAGCATGTGCAACATTTCTACAACTCCT	1938
Qy	2052	GACAGCCTCCATTGATGTATTCAGAACTGGGAGAAAAGATTCGCGGATTTTACTGATCT	2111
Db	1939	GACAGCCTCCATTGATGTATTCAGAACTGGGAGAAAAGATTCGCGGATTTTACTGATCT	1998
Qy	2112	CCCCAAAGAGATCAGACATTACTTTATTGAATCAGCCCTTTTGGAGCTGTTTGTCTCTCAG	2171
Db	1999	CCCCAAAGAGATCAGACATTACTTTATTGAATCAGCCCTTTTGGAGCTGTTTGTCTCTCAG	2058
Qy	2172	ACTTTTCCATCAGGTCACAACTGCTGAAGTAAAGTTTGTCTTTCGAATGCACTGTCTCCT	2231
Db	2059	ACTTTTCCATCAGGTCACAACTGCTGAAGTAAAGTTTGTCTTTCGAATGCACTGTCTCCT	2118
Qy	2232	GCATCGACTTCAGTGCCTTCGTGGATTGGGAGTGGCTCGACTCTATTAAAGACTTTTC	2291
Db	2119	GCATCGACTTCAGTGCCTTCGTGGATTGGGAGTGGCTCGACTCTATTAAAGACTTTTC	2178
Qy	2292	CTTAAATTTGCAGAGCCTGAACTTTGATATCCAGCCCTTAGCCCTGCTGTGACACTGAG	2351
Db	2179	CTTAAATTTGCAGAGCCTGAACTTTGATATCCAGCCCTTAGCCCTGCTGTGACACTGAG	2238
Qy	2352	CATGATCACAGAAAGACATGCGTTTAAAGAACCAAGAGAGTGCAGAGACTATGCAACAA	2411
Db	2239	CATGATCACAGAAAGACATGCGTTTAAAGAACCAAGAGAGTGCAGAGACTATGCAACAA	2298
Qy	2412	GATCAACAGCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCTGGAGCCCAACCGAGTC	2471
Db	2299	GATCAACAGCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCTGGAGCCCAACCGAGTC	2358
Qy	2472	CAAGGTCTCGGGTCCCTGGTGTAGAACTGAGGAAGATCTGCAACCTGGGCTCTCAGCGCAT	2531
Db	2359	CAAGGTCTCGGGTCCCTGGTGTAGAACTGAGGAAGATCTGCAACCTGGGCTCTCAGCGCAT	2418
Qy	2532	CTTCTACTGAGCTGGAAGACTTTGGTGTCTCCACCTTCATCATTTGACAAGACTCTTCCT	2591
Db	2419	CTTCTACTGAGCTGGAAGACTTTGGTGTCTCCACCTTCATCATTTGACAAGACTCTTCCT	2478
Qy	2592	GGACACCTTACTTTTAACTCAGAGACAGTGGAGCAGTGGAGCTGAGCTGCTCTCTAGCAC	2651
Db	2479	GGACACCTTACTTTTAACTCAGAGACAGTGGAGCAGTGGAGCTGAGCTGCTCTCTAGCAC	2538
Qy	2652	CTGCTTGCTACGCAGCAAGGATAGGTTTGGAAACCTATCATTTCTGCTCTTCTCTTAA	2711

Db	2539	CTGCTTGGCTAGCGAAGGAGTAGGTTGGAAAGCTATCATTTCTGTCCTTCTCTAA	2599
Qy	2712	GAGGAAAGCAGCTCTCTGTAGAAGCAAGACTTTCTTTTTTCTGGCTCTTTTCCCTTA	2771
Db	2599	GAGGAAAGCAGCTCTCTGTAGAAGCAAGACTTTCTTTTTTCTGGCTCTTTTCCCTTA	2658
Qy	2772	CAACTAAAGCCAGAAAACCTTGACAGAGTATGTGTGGGGTGTGTTTATATTTA	2827
Db	2659	CAACTAAAGCCAGAAAACCTTGACAGAGTATGTGTGGGGTGTGTTTATATTTA	2714
RESULT 7			
LOCUS	E10601	4400 bp	RNA linear PAT 29-SEP-1997
DEFINITION		cDNA encoding rat Nuclear Receptor 88(NUR88).	
ACCESSION	E10601		
VERSION	E10601.1	GI:22027657	
KEYWORDS	JP 1996023980-A/1.	(Norway rat)	
SOURCE	Rattus norvegicus		
ORGANISM	Rattus norvegicus		
REFERENCE			
AUTHORS	Ookura,N., Hijikuro,M., Miki,K. and Yamamoto,A.		
TITLE	NUCLEUS RECEPTOR FORMED IN APOPTOSIS OF CEREBRAL NEUROCYTE AND NUCLEUS RECEPTOR GENE CODING THE RECEPTOR		
JOURNAL	Patent: JP 1996023980-A 1 30-JAN-1996;		
COMMENT	TERUMO CORP		
	OS Rattus norvegicus (rat)		
	PN JP 1996023980-A/1		
	PD 30-JAN-1996		
	PF 15-JUL-1994 JP 1994164434		
	PI OOKURA NAGANARI, HIJIKURO MIDORI, MIKI KEIZABURO, PI YAMAMOTO AKIHIRO		
	PC C12N15/09.C07K14/705;		
	CC strandedness: Double;		
	CC topology: Linear;		
	CC hypothetical: No;		
	CC anti-sense: No;		
	CC Feature is identified by similarity;		
	FH Key Location/Qualifiers		
	FH source- 1..4400		
	FT /tissue_type='brain' /organism='Rattus norvegicus' FT		
	FT CDS 700..2586		
	FT /product='rat Nuclear Receptor 88(NUR88)' FT		
	FT misc_RNA 1576..1836		
	FT /product='DNA binding site'.		
FEATURES			
source	Location/Qualifiers		
	1..4400		
	/organism="Rattus norvegicus"		
	/mol_type="genomic RNA"		
	/db_xref="taxon:10116"		
ORIGIN			
Query Match	65.9%;	Score 2500;	DB 6; Length 4400;
Best Local Similarity	83.4%;	Pred. No. 0;	
Matches 3131;	Conservative 0;	Mismatches 535;	Indels 90; Gaps 22;
Qy	46	CGGAGTCTCTGCTCCCGCCCCCCCCCTCCAGCTCTGCTCTCTCCGCTCCCAT	105
Db	1	CCGAGTCTCTGCTCTCCCGCCCCCCCCCTCCAGCTCTGCTCTCTCCGCTCCCAT	60
Qy	106	ACACAGACGGGCTCACACCGGCTCCCTCACTTCGCACACACAGACACAAAGCGCGCACACAG	165
Db	61	ACACAGACGGTCAACACCGCTCTTCACTTCGCACACACAGACACACGGCGGCTCACAC	120
Qy	166	GCTCCGACACACACT--TGCTCTCCGGCGGCGGTACACCCCTCTTGGCGGTAGGCCCTT	223
Db	121	GCTCCGACACACACTCACTCTCCCGGGCGGTACACCCCTCTCTCTCGGCGGCTC	180

224 GCGGT- ---GCAGCGCGCGCGAGCTGGAGCGCCCTCCGCGGCTCACTTGCACG 278
181 GCGGTGTCGCGCGCGCGCGCGCGAGCGCCCTCAGGGCTCACTTGAACG 240
279 CTGACGGTCCGCGAGTGGCGGTGGAGGTGGAAACAGCGCGCGGCACTCCCTCCCTGGTC 338
241 CTGACAGAGCGGCGAGTGGCGGTGGAGGTGGAAACAGTGGAGACATCTTAGCCCCCTGGTC 300
339 ACAGCCCAAGCAGAGCGCGCGGGAACCTCTCGGCTGTGCTCTCCCATAGTTCGGGATC 398
301 GCAGCGGAGAGCTGGAGC-CTGCGGAACCTCTCGCGCGGCTCTCCCATAGTTCGGATC 359
399 GCAGANTCCCCACAGCC- ---GCTCACCGCTCTCGGAGCGCTGGGGCTGTACACGC 455
360 GCAGANTCCCGAGCGCGCTGCTCACCGCTCTGGAGCGCTGGGTTTGTGACCGC 419
456 AGCCCTTCGCGAGCAGCTGTGACTCCCCCGCCAGTGCAGATTTTCGGGACAGCTCTCTA 515
420 AGCCCTTCGCGAGCAGCTGTGACTCTCCCCAATCCAGATTTTCGGGCTGCTCTCTA 479
516 GAAACTCGCTCTAAGACGGAACCGGCAAGCACTCAAGGCCACTGGGGAAGAGGGAG 575
480 GAAACTCGCTCTAAGACGGAACCTCCAGAAACCCAAAGCCACTGCGGAGAGCGCAG 539
576 CCGGCAAGCCCGGCGCTGAGCCTGGAGCCCTTAGCGTGGCGGCGCA- ---CTGC 629
540 CCGGCAAGCCCGGCGCTGAGCCTGGAGCCCTCAAGAGCGGCGCAGACAGCGCGCG 599
630 CCGGCTTCGCTCGCGGAGCTGCTCTCTACTACTCTCAGCCTCGCTGGAGAGAC 689
600 GACTGCTTCGCTATCCGAGCTCCCGCTCTCTACTCTCAGCTTCGCTGGAGAGAC 659
690 CCGGAGCCCACTTCAGCGGCGAAGATACCTTCAGATATGCCCTGGCTCAAGCCCA 749
660 CCGGAGCCCACTTCAGCGGCGAAGATACCTTCAGATATGCCCTGGCTCAAGCCCA 719
750 ATATAGCCCTTCCTCCAGGTTCCAGTTATGGGCGGAGACATACAGCTCGGATAC 809
720 ATATAGCCCTTCGCTTCGCGGCTGCACTTATGCGAGCGAGATATGGCTCGGAATAC 779
810 CAGGAGATCATGAACCCCGACTACACCAAGCTGACCTGAGCAGCTGGAGACAT 869
780 CACAGNAATCATGAACCCCGACTATGCCAAGCTGACCATGGACCTCGGTAGCAGGGAT 839
870 CAGGCTACAGCACCATCTCCTGCGGAGCATAGTACCTTCGGAGGGCTACTCGAG 929
840 CATGGCCACCGGCAAGCATCTGCGGAGCTTCAGTACCTTCATGAGGGCTACCCAG 899
930 GAATACGAACTCAAGCTTCCTGCTGTACCAATGCA- ---CGGCGCTT 977
900 CAGCTGCGAACTCAAGCCCTCTGCTGTACCAATGCGCTTCGCGGCTTCGCGCTT 959
978 GATCAAGTGGAGGAGGCGCGCGGCTACCATACCATACCATACCATACCATACCA 1037
960 GATCAAGTGGAGAGGGTTCGAGCATGCTTACCAACACCATACCATACCATACCA 1019
1038 CCACCAACCATACCATACCATACCATACCATACCATACCATACCATACCATACCA 1097
1020 TCATCACCAACCATACCATACCATACCATACCATACCATACCATACCATACCA 1073
1098 GGAGGAGAGGTGCTGCGGAGCGCTTCATGTACTTCAAGCAGTCCCGGCTACCGC 1157
1074 CGAGGAGAGGTGCTGCGGAGCGCTTCATGTACTTCAAGCAGTCCCGGCTACCGC 1133
1158 CACCAAGCGGCTTCGCGGAGCGGCGGCTTATGGGAGCGGAGCTGCGCTCGGC 1217
1134 GACCACTCAGGCTTCGCGGAGCGGCGGCTTATGGGAGCGGAGCTGCGCTCGGC 1193
1218 GCGGCTGCTGATCGCAACCGGCGCTGCTGGAACCGCGGATGAAGGCGGTCCCGCGT 1277
1194 GCTGCTGCTGCTGCTGCGGAGCGCTGCTGAGCCGCGAGATGAAGGCGAGTCCCGCA 1253
1278 GCGCGGCGCGGCTTCGCGCTCTTCACTTCAAGCCCTCGCGCGCGCATCCCCCGCGC 1337

1254 GCGGCTGCTGCGGCTTCGCGATCTTCTTCAAGCCCTCACGCGCACACCTCCCGCGC 1313
1338 CAGCCCGCGCGCGCGCGCTACCTCGGTACGACCGAGCGCGCTGCGCGCTCAGCCT 1397
1314 CAGCCAGCGCGCGCGCGCTACCTCGGTATGACCCCGCGCGAGCTGCGCTCAGTCT 1373
1398 GCGGCTGGAGCGCGAGCGCGCGCGCGAGCGCGCGCGCTTGAGAGCGCACCGTA 1457
1374 ACCCTGGAGCGCGCGCGCGCGCGCGCGAGCGCAAGCTGCTGCGCTCGAGGGCGCATCCGTA 1433
1458 CCGGCTGCGGCTGGCGCAAGAGCGCGCGCGCTGCGCTTCCCGCTCTCGGCTCACGCG 1517
1434 CCGGCTGCGGCTGGCGCAAGAGCGCGCGCGCTGCGCTTCCCTCGCTCGGCTCACAGC 1493
1518 CTCCCTTACCGGCTCGAGCTGCTGGGCGAGAGTCCAGCCTGCGCTGCGCGCGCGAG 1577
1494 GTCCCTTACCGGCTCGAGCTGCTGGAGAGAGCGCGCGCTACCATCGCCACCAATAG 1553
1578 GAGCTGCTGCTGCGGAGGCGAGCTGCTGCGCTGCGGCGGAGCAAGCGCGCTGCGAGCA 1637
1554 GAGCTCATCATCTCGGCGAGGCGAGCTGCTGCTGCTGCGGCGGAGCAATGCTGCGCGCAGCA 1613
1638 CTACGCGGTGCGAACTGCGAGGCTGCAAGGCTTTTTCAAGAGAACAGTGCAGAAAAA 1697
1614 CTACGAGTCCGCACTGCGAGGCTGCAAGGCTTTCTTCAAGAGAACCGTGCAGAAAAA 1673
1698 TGCAAAATATGTTGCTGCGCAAAATAAACTGCGGCGGAGCAATGCTGCGCGCGAG 1757
1674 CGCAAAATATGTTGCTGCGCAAAATAAACTGCGGCGGAGCAATGCTGCGCGCGAG 1733
1758 ATGTCAGTACTGCTGATTTCAAGAGTCTCAGTGTGGAATGTTTAAAGAGTGTGCTCG 1817
1734 ATGTCAGTACTGCTGATTTCAAGAGTCTCAGTGTGCGGATGTTGGAAGAGTGTGCG 1793
1818 TACAGATGCTGAAAGGAGGAGAGTGTGCTGCTTCCAAACCAAGAGCGCATTACA 1877
1794 TACAGATGCTGAAAGGAGGAGAGTGTGCTGCTTCCAAACCAAGAGCGCATTACA 1853
1878 ACAGGAACCTTCTCAGCCCTTCCACCTTCTCCTCAATCTGCATGATGAATGCCCTGT 1937
1854 ACAGGAGCCTCGCAGCGCTCCCGCGCTCCCGATCTGTATGATGAACGCGCTGT 1913
1938 CCGAGCTTTAAGAGCTCAACCGCAGAGATCTGATTTATTCAGATCTGCTCCAGTGA 1997
1914 CCGAGCTTTAAGAGCTCAACCGCAGAGACCTTGAATTTACTCCAGATCTGCTCCACCA 1973
1998 CCGAGCTGCTGAGGCGAGAGTGTGAGCATGCTGCAAACTTCAAACTCTCAGAGC 2057
1974 CAGGCGCTGCGGCGAGAGCTGAGCATGCTGAGCGAGTCTACACCTTCTGAGCGC 2033
2058 CTCCATTTGATGATCCAGAGCTGGGAGAGAGATTCGCGGATTTACTGATCTCCCGCA 2117
2034 CTCCATCGAGCTGCTCAGAGCTGGGAGAGAGATTCGCGGATTTACTGATCTCCCGCA 2093
2118 AGAAGTCAAGATTTATTTTGAATCAGCTTTTGGAGCTGTTTGTCTCAGATCTTC 2177
2094 AGAAGTCAAGATTTATTTAAGATCAGCTTTTGGAGCTGTTTGTCTTGAAGCTTTC 2153
2178 CATCAGGTCAACACTGCTGGAAGTAAAGTGTGCTTCAAGTGAATGCTGCTGATCG 2237
2154 TATCAGGTCAACACTGCTGGAAGTAAAGTGTGCTTCAAGTGAATGCTGCTGATCG 2213
2238 ACTTCAGTGCCTTCGTTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAA 2297
2214 ACTTCAGTGCCTTCGTTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAA 2273
2298 TTTGAGAGCTGAACCTTTGATTTCAAGCTTTAGCTGCTGCTGTCAGCATGAGCATGAT 2357
2274 TTTGAGAGCTGAACCTTTGATTTCAAGCTTTAGCTGCTGCTGTCAGCATGAGTATGAT 2333
2358 CACAGAAAGACATGGTTTAAAGAACCAAGAGTTCGAGAGCTATGCAAGATCAT 2417

Qy	3464	GATTACATGTAAC	TGTTTGGTTGTG	TCTATCAACCCACCAGAGT	CCCCAAACTTG	3522		
Db	3471	GAATGCATATGAC	TGTTTCCGTGCTGTG	TGTCACACCTGGA	CAGAGTCTCTCTAAACCTT	3530		
Qy	3523	CTTCAGTTATAGTAA	CTGACTGCTATAT	TCAATTCAGNAGC	-----GCCATAAGTCAGT	3575		
Db	3531	CTTTAGTTTGTAGCA	AGTTCCTGATTCCT	CAATTCAGAAAGCCCA	AGGAGCAATTCGGTGACT	3590		
Qy	3576	TGAGTATTTTCATC	CCCTAGATAAGAA	CATGCAAAATCAGCAGGA	ACT--GGTCATACAGGGTA	3634		
Db	3591	CGATCAAGGTTTAA	CCCTAGGAGAA	CATGCAATTAAGTAGGA	ACTGGTTCAGACAGGGTA	3650		
Qy	3635	AGCACAGGGACAA	TAAAGGATTTTAT	TAGATAT-----AA	TTAATTTTTCGTTATTTG	3687		
Db	3651	AGCACAGAGATGA	TAAAGGATTTAT	TATATAAATATATA	TATAAATAATTTTTCGTTATTTG	3710		
Qy	3688	TTAAGGAGCAAA	TTTGGAGAGCA	AGCAATCTTTT	3723			
Db	3711	TTA--TAGACA	ATTTTGGAAAGCA	AGAGATCATCT	3744			
RESULT 8								
AX587468		AX587468		4400 bp		DNA		
LOCUS		Sequence 3 from Patent WO0246391.		linear		PAT 10-JAN-2003		
DEFINITION		AX587468						
ACCESSION		AX587468						
VERSION		AX587468.1 GI:27656292						
KEYWORDS		Rattus norvegicus (Norway rat)						
SOURCE		Rattus norvegicus						
ORGANISM		Rattus norvegicus						
REFERENCE		1 Guenther,C. and Allen,K.D.						
AUTHORS		Transgenic mice containing nori gene disruptions						
TITLE		Patent: WO 0246391-A 3 13-JUN-2002;						
JOURNAL		Deltagen, Inc. (US)						
FEATURES		Location/Qualifiers						
source		1..4400						
		/organism="Rattus norvegicus"						
		/mol_type="unassigned DNA"						
		/db_xref="taxon:10116"						
ORIGIN								
Query Match		65.9%		Score 2500;		DB 6; Length 4400;		
Best Local Similarity		83.4%		Pred. No. 0;				
Matches 3131;		Conservative		0; Mismatches		535; Indels 90; Gaps 22;		
Qy	46	CGGAGTCTCTGCT	CCCGCCCCCACCCTCC	CAGACTCTGCTCTCT	CGGTCCTCCCAT	105		
Db	1	CGGAGTCTCTGCT	CCCGCCCCCACCCTCC	AGGCTGCTCTCTCT	CGGTCCTCCCAT	60		
Qy	106	ACACAGCGCGCTCA	CCCGCTCCCTCACT	TCGCACACACAGACAC	AGCGCGCACACAG	165		
Db	61	ACACAGACGCTCA	CAACCGCTCTCTCACT	TGCACACACAGACAC	ACGCGGCTCACAC	120		
Qy	166	GCTCCGCACACACT	--TGGCTCTCCGCGCGCT	CACACCCCTCTTGGCCCT	CATGAGCCCTT	223		
Db	121	GCTCCGCACACACT	CTCACTCTCTCCGCGCGCT	CACACCCCTCTCTCT	CGGCGCCCTC	180		
Qy	224	GCCGCT----	GCAGCGCGCGCGCG	AGCTGGACGCCCT	CTCCGGGCTCACTTTG	CAACG 278		
Db	181	GCCGCTGTGCGG	CGCGCGCGCGCG	AGCGGACGCCCTT	CCAGGGCTCACTTTG	CAACG 240		
Qy	279	CTACGCTGCGG	CAGTGGCGTGGAGT	GGGAAACAGCGCGG	GCATCTCTCCCTCCCT	CGGTC 338		
Db	241	CTGNACAGCGG	CAGTGGCGTGGAGT	GGGAAACGTGGCG	ACATCTTAGCCCCCT	CGTC 300		
Qy	339	ACAGCCAA	CGCAGGACCGCGG	GAACCTCTCGGCTGT	GCTCTCCCATGAGTC	GGGATC 398		
Db	301	GCAGCCGG	GACTGGAG--CTGG	GGAACTCTCGGCGCG	CTCTCCCATGAGTT	GGGATC 359		

QY 399 GCAGCATCCCCACACAGCC---GCTCACCGCCTCCGGGAGCGCTGGGCTTGTACACGC 455
DB 360 GCAGCATCCCCAGCAGCGCTGCTCACCGCCTCTGGGAGCGCTGGGTTGTGACCGC 419
QY 456 AGCCCTTCGGGACAGCAGCTGTGACTCCCGCCAGTGCAGAGATTTCCGGACAGCTCTCTA 515
DB 420 AGCCCTTCGGGACAGCAGCTGTGACTCTCCCGCAATCCAGATTTCCGGGCTCGCTCTCTA 479
QY 516 GAAACTCGCTCTAAGACAGGAACCGCACAGCACTCAAGGCCACTCAAGCCGAGAGGGCAG 575
DB 480 GAAACTCGCTCTAAGACAGGAACCTCCACAGAAACCAAGCCCACTTCGCGGAGAGCGCAG 539
QY 576 CCGCGCAAGCCGGGCGCTGAGCCTGAGCTGAGCCCTTAGCGGTGCCGGGACGA-----CTGC 629
DB 540 CCGGACAGCCCGGGCGCTGAGCCTTGAGCCCTCAACAGAGCGGGCCAGCACAGCGGCGGC 599
QY 630 CGCGCGCTTCGCTTCGCGGAGCGTCCGCTCTCTCTACACTCTCAGCCCTCCGCTGGAGAGAC 689
DB 600 GGCTGCTTCGCTATCCCGAGCGTCCCGGCTCTCTACACTCTCAGCCCTCCGCTGGAGAGAC 659
QY 690 CCCAGCCCCACATTCAGGGGCAAGATACCTCCAGATATGCCCTCGCTCCAGGCCA 749
DB 660 CCCAGCCCCACATTCAGGGGCAAGATACCTCCAGATATGCCCTCGCTCCAGGCCA 719
QY 750 ATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGCGCAGACATACAGCTCCGGAATACAC 809
DB 720 ATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGCGCAGACTTATGGCTTCGGAATACAC 779
QY 810 CACGGAGATCATGAACCCCGACTACACCAAGCTGACCATGGACCTTGGCAGCACTCAGAT 869
DB 780 CACAGAAATCATGAACCCCGACTATGCGAAGCTGACCATGGACCTCGTAGCACGGGAT 839
QY 870 CACGGCTACAGCCACACAGTCCCTGCGCCAGCATCAGTACCTTGGTGGAGGGCTACTCGAG 929
DB 840 CATGGCCAAGCCACAGCTCCCTGCGCCAGCTTCACTTCACTGAGAGGGCTTACCCAG 899
QY 930 CAACTACGAACCTCAAGCCTTCCTGCTGTACCAATGCGAG-----CGGCCCTT 977
DB 900 CAGCTGCGAACTCAAGCCCTTCCTGCTGTACCAAAATGCGCCCTTTCGGGCTTCGGGCTTT 959
QY 978 GATCAAAGTGGAGGAGGGCGGCGCCAGCTACCATCACTACCATCAACACACACCA 1037
DB 960 GATCAAGATGGAAGAGGGTGGGAGCATGCTTACCAACACACCATCACTACCATCA 1019
QY 1038 CCACCAACCACTACACAGCAGCAGCATCAGCAGCCATCCATTCCTCCAGCCTCCAGCCC 1097
DB 1020 TCATCAACCAACCAACAGCAG-----CAGCAGCGTCCATTCCTCCTCCTCGGCC 1073
QY 1098 GGAGGACGAGGTGCTGCCAGACCTCCATGTACTTCAAGCAGTCCCGCAGCTCCACCC 1157
DB 1074 CGAGGACGAGGTACTGCCAGCAGCACTCCATGTACTTCAAGCAGTCTCCGCGCTACGCC 1133
QY 1158 CACCAGCGGCTTCCCGCCGAGCGGGGGGTTATGGGAGGAGGCACTGCCCTCGGC 1217
DB 1134 GACCACTCAGGCTTCCCGCCGAGCGGGGGCTGTGGAGCAGAGCTGCCCTCTGC 1193
QY 1218 GCCCGCTGCATCGCACCGGCCCGCTGCTGGACCCCGCGATGAAGCGGTCCCCACGGT 1277
DB 1194 GCCTGCTGCATCGCTCCGGGACCGCTGCTGGACCCCGCAGATGAAGGAGTGCCCAAT 1253
QY 1278 GGCAGCGCGGCTTCCCGCTTCCATTTCAAGCCCTCGCGCGCGATCCCGCGGCC 1337
DB 1254 GGCAGCTGTGCGGCTTCCCGATCTTCAAGCCCTCACCGCCACACCCCTCCCGGCC 1313
QY 1338 CAGCGCGCGCGGCCACCACTCGGCTACGACCCGAGCGGCGCTGCCGCTCAGCCT 1397
DB 1314 CAGCGCGCGCGGCCACCACTCGGCTATGACCCCAAGCGCGAGCTGCGCTCAGTCT 1373
QY 1398 GCGCTGGGAGCGCAGCGCGCGGCCAGCAGCGCGCGCTTGGAGGCGCACCCGTA 1457
DB 1374 ACCCTTGGGAGCGCGCGCGGCCAGCGGAGCGAGCTGTGCTCGGCTCGAGGGCCATCCGTA 1433

QY 1458 CGSGCTGCGCTGGCCAAAGAGGGCGGCCCGCTGGCCTTCCCGCCTCTCGGCCTCAGCC 1517
DB 1434 CGSGCTGCCGCTGGCCAAAGAGGAGGCCACGTTGACTTCCCTCCGCTGGGCTCAGAGC 1493
QY 1518 CTCCCTTACCGGCTCGCAGCCTTCGTCGGCGAGAGTCCAGCCTGCGTCGCGCCGAGCAG 1577
DB 1494 GTCCCTTACCGGCTCGCAGCCTTCGTCGGGAGAGGCCCCAGCCTACCATCGCACCAATAG 1553
QY 1578 GAGCTGCTGCTGCGGAGGGGCACTGTGTCGTCGTCGGGGGCAACGCGCCTGCGCAGCA 1637
DB 1554 GAGCTCATCATCGCGGAGGGCACTGTGTCGTCGGGGGCAAACTGCTGCTGCCAGCA 1613
QY 1638 CTACGCGCTGCGAACTCGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGCAGAAAAA 1697
DB 1614 CTACGAGTCCGACCTGCGAGGCTGCAAGGGCTTCTCAAGAGAACCGTGCAGAAAAA 1673
QY 1698 TCGAAAAATATGTTGCTGCGCAAAATAAAAACTGCCAGTAGACAAGAGAGCTCGAAACCG 1757
DB 1674 CGCAAAATATGTTGCTTGGCAAAATAAAAACTGCCCGGTAGACAAGAGAGCTCGAAATCG 1733
QY 1758 ATGTCAGTACTGTCGATTTTCAGAAAGTCTCAGTGTGGAAATGGTAAAAAGATTGTCG 1817
DB 1734 ATGTCAGTACTGTCGATTTTCAGAAAGTCTCAGTGTGGGATGGTGAAGAACTGTCG 1793
QY 1818 TACAGATAGTCTGAAAGGAGGAGGAGTCTGCTGCTTCCAAACCAAGAGGCCATTACA 1877
DB 1794 TACAGATAGTCTGAAAGGAGGAGGAGTCTGCTGCTTCCAAACCAAGAGGCCACTACA 1853
QY 1878 ACAGAAACCTTCTCAGCCCTTCTCCACCTTCTCCTCAATCTGCATGATGAATGCCCTGT 1937
DB 1854 ACAGAGCCCTCGCAGCCCTCCCGCATCTCCTCGATCTGTATGATGAAACGCTTGT 1913
QY 1938 CCGAGCTTTAAACAGACTCAACACCCAGAGATCTGATTTATTCAGATATCTGTCCTCACTGA 1997
DB 1914 CCGAGCTTTAAACAGAGCGCAACCGCCAGAGACCTTGATTACTCCAGATCTGTCCTCCACGA 1973
QY 1998 CCGAGCTGCTGAGCAGCAGATGCTGAGCATGTGCAACAAATTTCTACAACCTCCTGACAGC 2057
DB 1974 CCGAGCCACTGCGGGGACAGACGCTGAGCAGCTGCGAGAGTCTTACAACCTTCTGAGCGC 2033
QY 2058 CTCCATTTGATGATTCAGAAAGCTGGGCGAGAAAGATTCGCGGATTTTACTGATCTCCCAA 2117
DB 2034 CTCCATCGAGCTGTCAGAAAGCTGGGCGAGAAAGATCCCGGATTCAGTATCTCCCAA 2093
QY 2118 AGAAGATCAGACATTAATTTGAATCAGCCTTTTGGAGCTGTTTGTCTCAGACTTTC 2177
DB 2094 AGAAGATCAGAGCTTACTTATAGAAATCAGCCTTTTGGAGCTGTTTGTCTTACAGCTTTC 2153
QY 2178 CATCAGGTCAACACCTGCTGAAGATTAAGTTTGTCTGCAATGGACTTGTCTGATCG 2237
DB 2154 TATCAGGTCAACACCTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTTGTCTGACCG 2213
QY 2238 ACTTCAGTCCCTTCGTTGGATTTGGGAGTGGGCTCGACTCTATTAAGAAGCTTTTCTTAA 2297
DB 2214 ACTTCAGTCCCTTCGCGGATTTGGGAGTGGGCTCGACTCCATTAAGAAGCTTTTCTTAA 2273
QY 2298 TTTGAGAGCTGAACTTTGATATCAAGCCTTAGCCTGCTGAGCAGTGAAGCATGAT 2357
DB 2274 TTTGAGAGCTGAACTTTGATATCAAGCCTTAGCCTGCTGAGCAGTGAAGTATGAT 2333
QY 2358 CACAGAAAGACATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAGATCAC 2417
DB 2334 CACAGAGCAGATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAGATCAC 2393
QY 2418 AAGAGTTTTAAAGACCAACAGAGTAAGGACAGGCTCTGGAGGCCACCGAGTCCAGGT 2477
DB 2394 AAGAGCTTTAAGAGACCAACAGAGAGGAGCAGGCTCTGGAGCCCTCAGAGCCAGGT 2453
QY 2478 CTTGGTGCCTTGGTAGAACTCAGGAAGATCTGCACCTTGGGCTCCAGCGCATCTTCTA 2537
DB 2454 CTTTCGCGCACTTGGTGAAGTCTGAGGAAGATCTGCACCGAGGCTCCAGCGTATCTTCTA 2513
QY 2538 CCTGAAGCTGGAAGACTTGGTGTCTCCACTTCCATCATTTGACAAAGCTCTTCTGGACAC 2597


```

Db      2514  |||||CTGAAGCTGGAGGACTTGGTGTCCCACTTCTGTCAATCGACAAGCTCTCTCTTGATAC 2573
Qy      2598  CCTACTCTTCTAATCAGGAGCAG-TGGAGCAGTGAAGTCCCTCTCTCTAGCAGCTGCT 2656
Db      2574  CTTGCTCTTCTGAGCAGGGAAGCCTGACAGAGAGCTACTTGTCTCTGCTGCACTGGTC 2633
Qy      2657  TCGTACGACGAAAGGATAGTGTGGAAACCTATCATTTCTGCTCTTCTTCTTAAGAGGA 2716
Db      2634  ATTAAGTCGAGCAAAAGGATGGTTTGAACACCT-GCCCTCTATCTCTCTCCAGGGAA 2692
Qy      2717  AAAGCAGCTCTGTAGAAAGCAAGACATTTCTTTTCTGCTCTCTTTCTCTTCAACCT 2776
Db      2693  AAAGCAGCTCCCATAGAAAGCAAGACATTT-TTTTCTCTGSCACCTTCTCTTCAAC 2751
Qy      2777  TAAAGCCAGAAACTTGCAGAGTATTGTGTGGGTGGTGTGTTTATATTATTTAGGCACTGG 2836
Db      2752  TAAAGCCAGAACTTGCAGAGTATTGTGTGGGTGGTGTGTTTATATTATTTAGGCTTGT 2811
Qy      2837  GGATGGGTGGAGGGGT--TATAGTTTCATGAGGGTGTTCCTAAGAAATTCCTAACAAG 2894
Db      2812  CGGTGGGTGGAGGGGTAAATAGTTTCATGAGGCTTTCTAAGAAATTCCTGACGAAG 2871
Qy      2895  CACTTTTGGACAATCTATCCAGCAG-----GAAAAAAGGATATATTAACCTGTTT 2948
Db      2872  CACTTTTGGATGATCTATCCAGCAGTGGGTGGGAGAAAGGATAATATACTGTTT 2931
Qy      2949  AAAA--CTCTTCTCGGGAATCAATATAGTTGTTTGTATTAAAGCAAGACAGCC 3006
Db      2932  AAAAATCTTTCGGGGGATATGACTATGTTGCTTTGTATTAAATTAAGAACAGCC 2991
Qy      3007  AAGGGTTG-TTCCGACGGTAGGATGTGTCTTAAAGATTGGTCTCTTGAATAATGCTTC 3065
Db      2992  AAGGGCTGTTTACCAGGCTAGGGCTGTGCTTAAAGACTGATCCCTTTAGTATGACTTC 3051
Qy      3066  CTGTATCAAGTAGTATGTGGTGCACAAACAGGAGAAA-----CTTCTCTTT 3113
Db      3052  CCGGATC-GAGGCACATAAGTGGTGCAATAGAGGGGGGAAATCTTCACTTCTTCATTT 3110
Qy      3114  TAAATCTCTTCTCTTATTATTAAATAATGATGAGGATGAGGATACCTTCAAAATCA 3173
Db      3111  CTTTCTCTCTCTTAAATAATGATGAGGATGAGGATGAGGATATCTCAAAATCA 3170
Qy      3174  GACATGGCAAAACAATAATGCTGTGTTCTTCAATAAACAAGTGCATTTTAAAGTGC 3233
Db      3171  GACTTAGCAAAATGATAATGGCTATTCGCTTCCACATCAAGTGCATTTTATAGGTGC 3230
Qy      3234  TGTCTTACTAAGTCTGTGTTTATTAATCTCTCTTATTTCTATATGAAATATAAAGAGGC 3293
Db      3231  TGTCTTACTAAGTCTGTGTTTGTGAATCTCTCTCTCTCTTATATGAAATATAAAGAGGC 3290
Qy      3294  AGTCATGTTAGCAATGACACGTTAATATCCCTAGCAGGGCTGTGTTCACTTCCCTGT 3353
Db      3291  AGTCATGTTATCAACAGGGGTGCTCATTTTCTAGCTACCTTGGTCCACTGCTCTGT 3350
Qy      3354  CGATCCCTCTGAGGTATGGCCATCCAGAGCTTTTAGGCCATTTCTGATGGAACACAGAT 3413
Db      3351  AGAACCTCTCGAGGTATGGCCCTTCTAAGACTTTTCAGGCCACTCTTGATGGAATTCGAC 3410
Qy      3414  CCC-----TGCCCTGACTGCTCCAGCTATCTGAAAGTGGATCAGATTATAAATCTG 3463
Db      3411  ACCCTCTCCCTCAACCCATGACTATCCAGATGCTCTGAAATGGGGATCAGGTTATAAATG 3470
Qy      3464  GATTACATGTAAGTCTGTTTGGTGTGT-TCTATCAACCCACAGAGTTCCTTAACTTTG 3522
Db      3471  GATTGCATGACTGTGTGCTGTGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3530
Qy      3523  CTTAGTTATAGTAATGACTGCTGGTATATTCATTCAGAAGC-----GCCATAAGTCACT 3575
Db      3531  CTTTAGTTGATAGCAAGTTCCTGATTCCTCCTCAATTCAGAAGCCCAAGGAGCATTTGGTGACT 3590
Qy      3576  TGAGTTATTTGATCCCTAGATAGACATGCAATCAGCAGGAACT-GGTCAATACAGGTA 3634

```

```

Db      3591  CGATCAAGGGTTAACCTTAGGAGAACATGCAAAATAGTAGGAACTGGGTTCAGACAGGTA 3650
Qy      3635  AGCACCAGGCAATTAAGGATTTTATAGATAT-----AATTTAAATTTTGTATTGCG 3687
Db      3651  AGCACCAGAGATGAAGATTTATATATATATATATATATATATATATATATATATATAT 3710
Qy      3688  TTAAGGACAAATTTTGGAGAGCAAGCAAAATCTTTT 3723
Db      3711  TTA--TAGACAATTTGGAAGCAAGAGAATCATCT 3744

RESULT 9
RATNORI
LOCUS   Rattus norvegicus mRNA for NOR-1, complete cds. linear ROD 08-FEB-1999
DEFINITION
ACCESSION D38530
VERSION   D38530.1 GI:1483194
KEYWORDS  NOR-1; neuron derived orphan receptor; thyroid/steroid receptor superfamily.
SOURCE    Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (sites)
AUTHORS   Ohkura,N., Hijikuro,M., Yamamoto,A. and Miki,K.
TITLE     Molecular cloning of a novel thyroid/steroid receptor superfamily gene from cultured rat neuronal cells
JOURNAL   Biochem. Biophys. Res. Commun. 205 (3), 1959-1965 (1994)
MEDLINE   95110348
PUBMED    7811288
REFERENCE 2 (bases 1 to 4400)
AUTHORS   Ohkura,N.
JOURNAL   Unpublished
REFERENCE 3 (bases 1 to 4400)
AUTHORS   Ohkura,N.
JOURNAL   Direct Submission
TITLE     Submitted (14-OCT-1994) Naganari Ohkura, National Cancer Center Research Institute, Growth Factor Division; 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan (E-mail:nohkura@ncc.go.jp, Tel:81-3-3542-2511(ex.4302), Fax:81-3-3542-8170)
COMMENT   On Aug 8, 1996 this sequence version replaced gi:643599.
FEATURES
            location/Qualifiers
            1..4400
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /strain="Wistar"
            /db_xref="taxon:10116"
            700..2586
            /note="thyroid/steroid receptor superfamily"
            /codon_start=1
            /product="NOR-1"
            /protein_id="BAA07535.1"
            /db_xref="GI:643600"
            /translation="MPCVQAOYSPSPGSGTATQYGYEYTEINPNVAKLMDLGS
            TGMATATSLPSTFSEMEGYPSSCLQYMPSPGPRPLKMEGRGEGHYHHH
            HHHHHHHHQOQSPSPGSGDEVLPSTSMYFKPSPPTPTTTPGPPQAGALWDD
            ELPAAGCIIAGPLDDPMKAAAPPIPKPSPHPHPAPSPAGHHGLYDPT
            AAALSLPLGAAAGSQAALGHPYGLPLAKSTATLTFFPLGLTASPTASSLLGES
            PSLSPNRRSSSGEGTCVCGDNACOHYGVRTCEGCKGFFKRTVQKNKYCLANK
            NCPVDRNRRCQYCFKQCLUSVGMVEVVRDLSLKGRRGLPSKPSLQOQPSOPS
            PPSPIKMMNALVRLATDPRDLDSYFCTDQATAGTDAEHVQOQFYNLLTASDVS
            RSWAEKIPGFDLPKEDOTLLIESAFLEFLVRLISRNTAEHKEFVNCGLVHRLQ
            LRSGEWLDSIKDFSLNLSLNDIOALACLALSMTIERHGLKPKVEELCNKITS
            SLKDHQRKQALPESEPKVLRLVELRLKICTQGLQRIFYKLKLELVSPSPVDKFLD
            TLPP"
            2587..>4400
            polyA_signal
            3126..3131
            polyA_signal
            4115..4120
            ORIGIN

Query Match      65.9%; Score 2500; DB 10; Length 4400;
Best Local Similarity 83.4%; Pred. No. 0;

```


2178 QY CATCAGGTCAACACATGCTGAAGATAAAGTTTGTGTTCTGCAATGAGCTTGTCTCTGATCG 2237
2154 Db TATCAGGTCAACACATGCTGAAGATAAAGTTTGTGTTCTGCAATGAGCTTGTCTCTGACCG 2213
2238 QY ACTTCAGTCCCTTCGTGGATTGTTGGGAGTGGCTCCGACTCTATTAAAGACTTTTCCCTTAA 2297
2214 Db ACTTCAGTCCCTTCGTGGATTGTTGGGAGTGGCTCCGACTCTATTAAAGACTTTTCTTTAA 2273
2298 QY TTTTGAGAGCCTGAACCTTTGATATCCAAAGCCTTAGCTTCCCTGTGAGCACTGAGCATGAT 2357
2274 Db TTTTGAGAGCCTGAACCTTTGATATCCAAAGCCTTAGCTTCCCTGTGAGCACTGAGCATGAT 2333
2358 QY CACAGAAAGACATGGTTTAAAGAACCAAGAGAGTGGAGAGCTATGCAACAAGATCAC 2417
2334 Db CACAGAGCGACATGGTTTAAAGAACCAAGAGAGTGGAGAGCTATGCAACAAGATCAC 2393
2418 QY AAGCAGTTTAAAGAACCAAGAGAGTGGAGAGCTATGCAACAAGATCAC 2477
2394 Db AAGCAGTTTAAAGAACCAAGAGAGTGGAGAGCTATGCAACAAGATCAC 2453
2478 QY CCTGGGTGCCCTGGTAGAACTGAGAAAGATCTGCACCCCTGGGCCCTCCAGCGCATCTTCTA 2537
2454 Db CCTTCGCGACCTGGTGGAACTGAGAAAGATCTGCACCCAGGCCCTCCAGCGTATCTTCTA 2513
2538 QY CTTGAAGCTGGAAGACTTGGTGTCTCCACTTTCATCAATGTAACAAGCTTCTTCTGGAAC 2597
2514 Db CTTGAAGCTGGAAGACTTGGTGTCTCCACTTTCATCAATGTAACAAGCTTCTTCTGATAC 2573
2598 QY CTTACTTCTTAATCAGGAGCAG-TGGAGCAGTGAAGCTGCTCTCTCTAGCACCTGCT 2656
2574 Db CTTGCTTCTGAGCAGGGAAGCCTGAGCAGAGCTACTTGTCTCTGCTGCACTGGTTC 2633
2657 QY TGTCTAGCAGCAAGGGATAGTGTGGAACCTCATCTTCTGCTCTTCTTCTTAAGAGGA 2716
2634 Db ATTAAGTGAGCAAAAGGATGGTGTGAACCTT-GCCCTCTATCTCTCTCCAGGGGA 2692
2717 QY AAAGCAGCTCTGTGAGAAAGCAAGACTTTCTTTTTTTTCTGGCTCTTTTCTTCAACACC 2776
2693 Db AAAGCAGCTCTTCCATGAAAGCAAGACTTTTCTTTTTTCTGGCACCTTTCTTCAACACC 2751
2777 QY TAAAGCCAGAAACTTGCAGAGTATGTTGGGGTGTGTTTATATTTAGGCAATGGG 2836
2752 Db TAAAGCCAGAAACTTGCAGAGTATGTTGGGGTGTGTTTATATTTAGGCTTTGT 2811
2837 QY GGATGGGTGGAGGGGT--TATAGTTTCATGAGGGTCTTCTTAAGAAATTTGTAACAAG 2894
2812 Db GGTGGCTGGAGGGGTAAATAGTTTCATGAGCTTTCTTAAGAAATTTGTCAGGAAG 2871
2895 QY CACTTTTGACATGCTATCCAGCAG-----GAAAAAAGGATAATATACTGTTTT 2948
2872 Db CACTTTTGATGATGCTATCCAGCAGTGGGTGGGGAAGAGGATAATATACTGTTTT 2931
2949 QY AAAA--CTCTTTCTGGGAATCCAAATATAGTTGCTTGTATTTAAACACAGACAGCC 3006
2932 Db AAAA--CTCTTTCTGGGAATCCAAATATAGTTGCTTGTATTTAAACACAGACAGCC 2991
3007 QY AAGGGTTG-TTCCAGGCTAGGATGCTGTTTAAAGATTGGTCCCTTGAAATATGCTTC 3065
2992 Db AAGGGCTGTTTACCAGGCTAGGCTGCTCTTAAGACTGATCCCTTAGTATGTAATTC 3051
3066 QY CTGTATCAAGGTAGTATGTTGGTGAACAAGAGCAGAAA-----CTTCTTTT 3113
3052 Db CCGGATC-GAGGCACATAAGTGGTGAACAATGAGGGGGGAAATCTTCAATTTCTTCAAT 3110
3114 QY TAAATTCCTCTCTCTTTATTTTAAACAAATGGTGAAGATGAGGATTAACCTTACAAATCA 3173
3111 Db TTTTCTCTCTCTCTTTTAAATGGAATGCAAAAAAAGATGGAATTAATCTTACAAATCA 3170
3174 QY GACATGGCAAAACAATAATGGCTGTTTGTCTCCATAAACAAGTCAATTTTAAAGTGC 3233
3171 Db GACTTAGCAAAATGATAATGGCTATTCGCTTCCACATACAGTCAATTTTATAGTGC 3230
3234 QY TGTCTTACTAGTCTTGTATTATTAATCTCTCTTTTATCTATATGGAATTAAGAGGC 3293

3231 Db TGTCTTACTAGTCTTGTGTTGTAAGTCTCTCTCTTATATGAAAAATAAGAGAGGC 3290
3294 QY AGTCATGTTAGCAATGACACGTTAATATCCTACAGAGGCTGTGTTCACTTCCCTGCT 3353
3291 Db AGTCATGTTTCAAAAGCGGTGCTCATTTTCTAGCTCACCTTGGTCCACCTGCTGCT 3350
3354 QY CGATCCCTTCTGAGGTATGGCCCATCCAAAGACTTTTAGGCCATTTCTGATGGAACCATG 3413
3351 Db AGAACCTTTCGAGGTATGGCCCTTCTAAGACTTTTCAGGCCACTCTTGTATGAAATTCGAC 3410
3414 QY CCC-----TGCCCTGAGCTTCCAGCTATCTCTGAAAGTGGATCAGATTTATAAATG 3463
3411 Db ACCCTCCCTCAACCCATGACTATCCAGATGCTCTGAAATGGGATCAGGTTTATAAATG 3470
3464 QY GATTACATGTAACATGTTTGTGTTGTTGTT-TCTATCAACCCACAGAGTTCCTTAACTTG 3522
3471 Db GATTGATATGACTGTGTTGCTGTGTTGTTGTTCACTGGAACAGAGTTCCTTAACTTT 3530
3523 QY CTTCAAGTTATAGTAACACTGACTGGTATATTTCAATTCAGAAAGC-----GCCATAAGTCA 3575
3531 Db CTTTAGTTGTAAGATTTCTCTGATTTCTTCCATTCAGAAAGCCCAAGGAGCATTTGGGTGACT 3590
3576 QY TGAGTATTTGATCCTTAGATTAAGAACATGCAAAATCAGCAGGAAT-GGTCATACAGGTA 3634
3591 Db CGATCAAGGGTTAAACCTTAGGAGAACATGCAAAATAGTAGGAACCTGGGTCCAGACAGGTA 3650
3635 QY AGCACCCAGGACAATAAGGATTTTATAGATAT-----AATTTAAATTTTGTATTGG 3687
3651 Db AGCACAGAGATGATAGGATTTTATATAATATATAATATAATTAATTTTGTATTGG 3710
3688 QY TTAAGGAGACAATTTTGGAGAGCAAGCAAAATCTTTT 3723
3711 Db TTA--TAGACAATTTTGAAGCAAGAGATCATCT 3744

RESULT 10
BC068150
LOCUS
DEFINITION
Mus musculus nuclear receptor subfamily 4, group A, member 3, mRNA
(CDNA clone IMAGE:30534822), partial cds.
ACCESSION
BC068150
VERSION
BC068150.1 GI:45829690
KEYWORDS
Mus musculus
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 5886)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
PUBMED
12477932
REFERENCE
2 (bases 1 to 5886)
Straussberg, R.


```

ACCESSION AJ011767
VERSION AJ011767.1 GI:3821742
KEYWORDS neuron-derived orphan receptor-1 alfa; NOR-1 gene; transcription factor.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 Cases-Langhoff, C., Castello, A., Martinez-Gonzalez, J. and Badimon, L. De-differentiated smooth muscle cells in atherosclerotic plaques express NOR-1
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 5115)
TITLE Castello, A.
JOURNAL Direct Submission
AUTHORS Submitted (08-OCT-1998) Castello A., Patologia Molecular y Terapeutica, Instituto de Investigaciones Biomedicas, C.S.I.C., Jordi Girona, 18-26, Barcelona, 08034, SPAIN
FEATURES
    source
        1..5115
            /organism="Sus scrofa"
            /mol_type="mRNA"
            /db_xref="taxon:9823"
            /cell_type="activated smooth muscle"
            /tissue_lib="smooth muscle"
        1..5115
            /gene="NOR-1"
            /db_xref="GI:3821743"
            /db_xref="GOA:O97726"
            /translation="MPCVQAYSPSPGSSYAAQTYGSYYTTEIMNPDYTKLMDLGS
            TRITATATSLPSFSTFMGYSNVELKPSCLYQMPSPRPLIKMEGRARHGHHH
            HDHHHHHQOQOQOQPPQOQPSIPSPSPPEDEVLPSLMVFKQSPSPPTTPVF
            PQAGLWEDALPSQACCIAPGLDPPKAVPTVAGARFPLFKPSPHPDPSPA
            GHHLYDPTAAALGLPIGAAAAAAGQAALGHPYGLPLAKRAALAFSPL
            GTTPTTSLGESSPSLSPSPNSTASGETGCAVGDNAACQHYGVRTCEGCKFFK
            RTVQNAKYVLANCPVDKRRNRQYCRFOKCLISVGMVKEVVRTDSLKRRRLP
            SKPKPLQEPSPQSPSPSPVCMNVALVRALTDSTPRDLDYSKYCPADQAAAGTAEH
            VQOFYNLTASIDVRSWAELKIPGTDLPKEDQTLLESFLEFLVRLSIRSNTAEH
            KVFNGLVHLRLQLRGFEWLDSIKDFSLRLQSLNLDIQALACLALSALSMITERHGL
            KSPKRVLECNKITSLSKEHSGKQALEPTEPKVLRALVELRKICTLGLQIFLYLKE
            DLVSPSPSIDKLFLDLPF"
        2044..5115
            /gene="NOR-1"
            /note="alfa transcript"
        2168..2173
            /gene="NOR-1"
            /note="alfa transcript"
        3545..3550
            /gene="NOR-1"
            /note="alfa transcript"
    polyA_signal
        2168..2173
            /gene="NOR-1"
            /note="alfa transcript"
    polyA_signal
        3545..3550
            /gene="NOR-1"
            /note="alfa transcript"
    ORIGIN
        Query Match 62.5%; Score 2371; DB 4; Length 5115;
        Best Local Similarity 87.4%; Pred. No. 0;
        Matches 2832; Conservative 0; Mismatches 295; Indels 114; Gaps 17;
        619 GGCAGACATGCGCGGCTTCGCTCGCGGACGCTCGCTCCTCCTACCTCAGCTCC 678
        Db 1 GGCAGCGCGCGCGCTTCGCTCGCGGACGCTCGCTCCTCCTACCTCAGCTCC 60
        679 GCTGGAGAGACCCCGGACCCCATTCAGCGCGGAGATACCTCCAGATATGCGCTGC 738

```

```

61 GCTGGAGAGACCCCGGACCCCATTCAGCGCGGAGATACCTCCAGATATGCGCTGC 120
739 GTTCAAGCCCAATATAGCCCTTCCTCCAGTTCAGTTATGCGGCGGAGACATACAGC 798
121 GTTCAAGCCCAATATAGCCCTTCCTCCAGTTCAGTTATGCGGCGGAGACATATGCG 180
799 TCGGAATACACCAAGAGATATGAACCCCGACTACCAAGCTGACCAAGCTGACCACTGGC 858
181 TCGGAATACACCAAGAGATATGAATCTGACTACCAAGCTGACCAAGCTGACCACTGGC 240
859 AGCACTGAGATACCGGCTACAGCCACCACTGCTCCCTGCGGAGATCAGTTCGTTGGAG 918
241 AGCAAGGAGATACAGCCACTGCTCAAGCTCCCTGCGGAGCTTCACTACCTTCATGGAG 300
919 GGTACTCGAGCAACTACGAACCTTCTCTGGTGTACCAATGAG----- 969
301 GGTACTTCAGCAACTACGAACCTTCTCTGGTGTACCAATGAGCGGCGG 360
970 ---CGGCCCTTCATCAAGTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1026
361 CTTGGCGCTTATCAAGTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
1027 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1065
421 GACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
1066 ---CAGCAGCCTTCACTTCCTCAGCTTCAGCGCGGAGGAGAGTGTGCGGCGGAGCACC 1122
481 CAGCAGCAGCCTTCACTTCCTCAGCTTCAGCGCGGAGGAGAGTGTGCGGCGGAGCACC 540
1123 TCCATGTACTTCAAGCAGTCCCAACCGTCCACCCACCAACCAACCAACCAACCAACCAAC 1182
541 TCTATGTACTTCAAGCAGTCCCAACCGTCCACCCACCAACCAACCAACCAACCAACCAAC 600
1183 GCGGCGGCGCTTATGGGACAGGACCTGCGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1242
601 GCGGCGGCGCTGTTGGGAGAGCGGCTGCGCTTCGCGGAGGCTGCATCGCGCGGCGGCGG 660
1243 CTGCTGGAGCCCGCGGATGAAGCGGTCCTCCAGTGGCGGCGGCGGCGGCGGCGGCGGCTTC 1302
661 CTGCTCGACCCCGGATGAAGCGGTCCTCCAGTGGCGGCGGCGGCGGCGGCGGCGGCTTC 720
1303 CACTTCAGCCCTCGCGCGGCGGATCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1362
721 CACTTCAGCCCTCGCGCGGCGGATCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
1363 GGTCTACGACCCGCGGCGGCTGCGCGGCTGCGCTGCGCTGCGGCGGCGGCGGCGGCGGCGG 1410
781 GGATAGACCCGCGGCGGCTGCGCGGCTGCGCGGCTGCGCGCTAGGAGCGCGGCGGCGGCGG 840
1411 ---GAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1467
841 GCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
1468 CTGCGCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1527
901 CTGCGCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
1528 GCGTCCAGCTTGTGGGCGAGAGTCCAGCCTGCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGG 1587
961 AGTCCAGCTTGTGGGCGAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
1588 TTTGGCGAGGCGAGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1647
1021 TCAGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
1648 CGAATCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1707
1081 CGCAGCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
1708 GTTTCGCTGGCAAAATAAAATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1767
1141 GTTTCGCTGGCAAAATAAAATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1900

```


QY	1768	TGTCGATTTT	CAGAAGTGCTCAGTGTGGAAATGGTAAAGAAAGTTGTCGGTACAGATAGT	1827
DB	1201	TGTCGATTTT	CAGAAGTGCTCAGTGTGGAAATGGTAAAGAAAGTTGTCGGTACAGATAGT	1260
QY	1828	CTGAAAGGGAGGAGAGGTCGTCTGCCTTCCAAACCAAAAGAGCCCATTTACAAACAGGAACCT	1887	
DB	1261	CTGAAAGGGAGGAGAGGTCGTCTGCCTTCCAAACCAAAAGAGCCCGTTACAGCAGGAACCC	1320	
QY	1888	TCTCAGCCCTCTCCACCTTCTCCATCGCATGATGAATGCCCTTGTCGGAGCTTTA	1947	
DB	1321	TCTCAGCCCTCTCCACCTTCTCCATCGCATGATGAATGCCCTTGTCGGAGCTTTA	1380	
QY	1948	ACAGACTCAACACCCACAGAGATCTTGATTTATTCAGATACCTGCCACTGACCAGGCTGCT	2007	
DB	1381	ACAGACTCAACACCCACAGAGATCTTGATTTATTCAGATACCTGCCCGCTGACAGGCCGCT	1440	
QY	2008	GCAGGCACAGATGCTGAGCATGTGCAACAAATCTTACAACTCCTCGACAGAGCTCCATTTGAT	2067	
DB	1441	GCAGGCACAGATGCTGAGCATGTGACACAGTCTTACAACTCTTCAACAGCTCCATTTGAC	1500	
QY	2068	GTATCCAGAGCTGGGCAGAAAAGATTCGGGATTTACTGATCTCCCCAAAGAGATPAG	2127	
DB	1501	GTATCCAGAGCTGGGCAGAAAAGATTCCTGGAATTTACTGATCTCCCTCAAGAGATCAG	1560	
QY	2128	ACATTACTTATTTGAATCAGCCCTTTTGGAGCTGTTGCTCCTCAGACTTTCCATCAGGTCA	2187	
DB	1561	ACATTACTTATTTGAATCAGCCCTTTTGGAGCTGTTGCTCCTCAGACTTTCCATCAGGTCA	1620	
QY	2188	AACACTGCTGAAGATAAGTTGTGTTCTGCAATGGAATGTTGCTCTGCACTCGACTTCAGTGC	2247	
DB	1621	AACACTGCTGAAGATAAGTTGTGTTCTGCAATGGAATGTTGCTCTGCACTCGACTTCAGTGC	1680	
QY	2248	CTTCGTGATTTGGGAGTGCTCGACTTATTTAAAGACTTTTCTTAAATTTTGCAGAGC	2307	
DB	1681	CTTCGTGATTTGGGAGTGCTCGACTTATTTAAAGACTTTTCTTAAAGACTTTTGCAGAGC	1740	
QY	2308	CTGAACCTTTGATATCCAAAGCTTAGCCCTGCTGCTGAGCATGTGAGCATGATCAAGAAAGA	2367	
DB	1741	CTGAACCTTTGATATCCAAAGCTTAGCCCTGCTGAGCATGTGAGCATGATCAAGAAAGA	1800	
QY	2368	CATGGGTTAAAGAACCAAGAGAGTGCAGAGCTATGCAACAGATCACAAGCAGTTTA	2427	
DB	1801	CATGGGTTAAAGAACCAAGAGAGTGCAGAGCTATGCAACAGATCACAAGCAGCTTA	1860	
QY	2428	AAAGACACCAAGATTAAGGACAGGCTCTGAGAGCCACCGAGTCCCAAGTCTTGGGTGCC	2487	
DB	1861	AAAGACACCAAGATTAAGGACAGGCTCTGAGAGCCACCGAGTCCCAAGTCTTGGGTGCC	1920	
QY	2488	CTGGTGAACCTGAGGAAGATCTGCACCTCTGGGCTCCAGGCGCATCTTCTACCTGGAAGCTG	2547	
DB	1921	CTGGTGAACCTGAGGAAGATCTGCACCTCTGGGCTCCAGGCGCATCTTCTACCTAAGCTG	1980	
QY	2548	GAAGACTTGGTGTCTCCACCTTCCATCATTTGAAGCTTCTTCTGGAACACCCCTACCTTTC	2607	
DB	1981	GAAGACTTGGTGTCTCCACCTTCCATCATTTGAAGCTTCTTCTGGAACACCCCTACCTTTC	2040	
QY	2608	TAATCAGAGCAG--TGAGGAGTGAGCTGCTCTCTCTCTAGCACCTGCTTGCTACCCAG	2666	
DB	2041	TGAGCAGAGCAGCCTGAGCAGGAGCTGCTTCACTCTGCTAGCAGCCACTTGTCTAAGCGG	2100	
QY	2667	CAAAAGGATAGTTTGGAAACCTATCATTTTCTCTCTCTTAAAGAGGAAAAGCAGCTC	2726	
DB	2101	CAAAAGGATAGTTTGGAAACCTATCATTTTCTCTCTCTTAAAGAGGAAAAGCAGCTC	2160	
QY	2727	CTGTAGAAAGCAAGAC-----TTTCTTTTTTTTTTCTGGCTCT	2763	

RESULT 12	HSNUCREC	LOCUS	DEFINITION
-----------	----------	-------	------------

2546 bp mRNA linear PRI 22-JAN-1996
ear receptor.

[illegible]


```
Db 601 |CTCCGCGCCAGTCCAGCCGCGGCGCCACCACTCGGCTATACCCCAACGCGCGAGCT| 660
Qy 1387 |GGGCTCAGCTCGCGCTGGAGCGCGAGCGCGCGGGGAGAGCCAGCGCGCGGCTTGAG| 1446
Db 661 |GCACCTAGTCTGCGCCCTGGAGCGCGCGCGCGAGAGGAGCAAGCTGCTGCGCTCGAG| 720
Qy 1447 |AGCCACCCGCTAGCGGCTGCGCTGGCCAAAGAGGGGGCGCCCGCTGCGCTTCCCGCTCTC| 1506
Db 721 |GGCCACCCATACGGGCTCGGCTGGCCAAAGAGGAGCGGCCACGCTGACCTTCCCTCGGTG| 780
Qy 1507 |GGCTCAGCCCTCCCTACCGGCTCAGCCTGCTGGGCGAGAGTCCAGCCTGCGGTGCG| 1566
Db 781 |GGCCTCAGAGCCTCCGCCACCGCGTCCAGCCTGCTGGAGAGAGCCCAACGCTCCCATCG| 840
Qy 1567 |CGCCCGCAGCAGGAGCTCGCTGCTGGCGAGGCGACGTGTCGCTGGGGGAGAGAGCGCC| 1626
Db 841 |CCACCCATAGGAGCTCATCTCTGGGAGAGGACATGTCGCGTGTGGCGGAGAGAGCGT| 900
Qy 1627 |GCTGCGAGCACTACGGCTGCGGAACCTGCGAGGGGCTGCAAGGGCTTTTTCAAGAGACA| 1686
Db 901 |GCTGCGAGCACTACGGAGTCCGACCTGCGAGGGCTGCAAGGGCTTCTTCAAGAGAGCG| 960
Qy 1687 |GTGCGAGAAATGCAAAATATGTTTGCCTGGCAATATAAACTGCCAGTAGAGAGAGA| 1746
Db 961 |GTGCGAGAAATGCAAAATATGTTTGCCTGGCAATATAAACTGCCAGTAGAGAGAGA| 1020
Qy 1747 |CCTCGAAACCGATGTCAGTACTGCTGATTTAGAAAGTGTCTCAGTGTGGAGTGTAAAA| 1806
Db 1021 |CCCGGAAACCGATGTCAGTACTGCGAGATTTAGAAAGTGTCTCAGTGTGGAGTGTAAAG| 1080
Qy 1807 |GAAGTTGTCCGTCAGATAGTCTGAAAGGGAGAGAGTGTCTGCTGCTTCCAAACCAAG| 1866
Db 1081 |GAAGTTGTCCGTCAGATAGTCTGAAAGGGAGAGAGTGTCTGCTGCTTCCAAACCAAG| 1140
Qy 1867 |AGCCATACAGAGAACCTTCTCAGGCTTCTCAGCCTTCTCCTTCCATCTGCAATGATG| 1926
Db 1141 |AGCCATACAGAGAACCTTCTCAGGCTTCTCAGCCTTCTCCTTCCATCTGCAATGATG| 1200
Qy 1927 |AATGCGCTTGTCCGAGCTTAAACAGACTCAACACCCAGAGATCTTGTATTTCCAGATAC| 1986
Db 1201 |AATGCGCTTGTCCGAGCTTAAACAGATGCAACACCCAGAGATCTTGTATTTCCAGATAC| 1260
Qy 1987 |TGTCCTACTGACAGGCTGCTGAGGCAAGATGCTGAGCATGTGCAACAATTTTACAAAC| 2046
Db 1261 |TGTCCTACTGACAGGCTGCTGAGGCAAGATGCTGAGCATGTGCAACAAGTTTCTACAAAC| 1320
Qy 2047 |CTCCTGACAGCTTCAATGATGATCCAGAGCTGGGAGAGAGAGATTTCCGGGATTTACT| 2106
Db 1321 |CTCCTGACAGCTTCAATGATGATCCAGAGCTGGGAGAGAGAGATTTCCAGGATTTACT| 1380
Qy 2107 |GATCTCCCCAAGAGATCAGACATTAATTTGAAATCAGCCTTTTGGAGCTGTTTCTGTC| 2166
Db 1381 |GATCTCCCCAAGAGATCAGACATTAATTTGAAATCAGCCTTTTGGAGCTGTTTCTGTC| 1440
Qy 2167 |CTCAGATTTTCCATCAGGCTCAAACTGCTGAAGATAAGTTTGTGTTCTGCAATGAGCTT| 2226
Db 1441 |CTCAGATTTTCCATCAGGCTCAAACTGCTGAAGATAAGTTTGTGTTCTGCAATGAGCTT| 1500
Qy 2227 |GTCTGTCATCGACTTTCAGTCTTCTGATTTTGGAGTGGCTGCTGCTCTATTAAGAC| 2286
Db 1501 |GTCTGTCATCGACTTTCAGTCTTCTGATTTTGGAGTGGCTGCTGCTCTATTAAGAC| 1560
Qy 2287 |TTTTCTTTAAATTTTCAGAGCTGCAACCTTGTATATCCAAAGCTTACGCTGCTGTGAGCA| 2346
Db 1561 |TTTTCTTTAAATTTTCAGAGCTGCAACCTTGTATATCCAAAGCTTACGCTGCTGTGAGCA| 1620
Qy 2347 |CTGAGCATGATCAGAGAGAGATCGGTTAAAGAACCAAGAGAGAGTCCAGAGCATATGC| 2406
Db 1621 |CTGAGTATGATCAGAGAGAGATCGGTTAAAGAACCAAGAGAGAGTCCAGAGCATATGC| 1680
Qy 2407 |AACAGATCAGAGCAGTTTAAAGAACCAACAGAGTAAAGGAGAGAGCTCTGAGCCACCC| 2466
```

```
Db 1681 |ACCAAGATCAACAAGCAGCTTAAAGACCAACAGAGAGGAGCAGGCTCTGGAGCCCTCG| 1740
Qy 2467 |GAGTCCAAAGGTCCTGGGTGCGCTGTAGAACTCTGAGAAAGATCTGACCTGGGCTCCAG| 2526
Db 1741 |GAGCTTAAAGGTCCTGGCGGCTGTGTAAGACTGAGAAAGATCTGTACCAGGGGCTCCAG| 1800
Qy 2527 |CGCATCTTCTTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCTTACCAAGCTC| 2586
Db 1801 |CGCATCTTCTTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCTTACCAAGCTC| 1860
Qy 2587 |TTCTTGGACACCTTACCTTTCTAA| 2610
Db 1861 |TTCTTGGACACCTTACCTTTCTGA| 1884
```

```
RESULT 15
AL358937
LOCUS
DEFINITION
Human DNA sequence from clone RP11-6013 on chromosome 9, complete
sequence.
ACCESSION
AL358937.16 GI:16973850
VERSION
AL358937.16
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 146889)
Pearce,A.
AUTHORS
Direct Submission
TITLE
Submitted (16-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
JOURNAL
On Nov 17, 2001 this sequence version replaced gi.16304457.
COMMENT
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Swi,
SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-6013 is from the library RPCI-11.1 constructed by the group of
Piet de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
```

FEATURES

```
Location/Qualifiers
1..146889
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
```

/clone="Rp11-6013"
/clone_lib="RPCI-11.1"

ORIGIN

Query Match 37.8%; Score 1435; DB 9; Length 146889;
Best Local Similarity 100.0%; Pred. No. 1.8e-223;
Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2360	CAGAAAGACATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAGATCACAA	2419
DB	25495	CAGAAAGACATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAGATCACAA	2554
QY	2420	GCAGTTTAAAGACACCAAGAGTGAAGAGCTCTGAGAGCCACCGAGTCCAAGGTCC	2479
DB	25555	GCAGTTTAAAGACACCAAGAGTGAAGAGCTCTGAGAGCCACCGAGTCCAAGGTCC	25614
QY	2480	TGGGTGCCCTGTGAGAACTGAGGAAAGATCTGACACCTGGGCTCCAGCGCATCTTCTACC	2539
DB	25615	TGGGTGCCCTGTGAGAACTGAGGAAAGATCTGACACCTGGGCTCCAGCGCATCTTCTACC	25674
QY	2540	TGAAGCTGAAGACTTGGTGTCTCCACCTTCCATCTTGAAGAGCTCTTCTGGACACC	2599
DB	25675	TGAAGCTGAAGACTTGGTGTCTCCACCTTCCATCTTGAAGAGCTCTTCTGGACACC	25734
QY	2600	TACCTTTCTAATCAGGACAGTGGAGCTGAGCTGCCCTCTCTCTAGCACCTGCTTGC	2659
DB	25735	TACCTTTCTAATCAGGACAGTGGAGCTGAGCTGCCCTCTCTCTAGCACCTGCTTGC	25794
QY	2660	TACGACGAAAGGATAGGTTTGGAAACCTATCATTTCTGCTCTTCTTAAAGAGGAAA	2719
DB	25795	TACGACGAAAGGATAGGTTTGGAAACCTATCATTTCTGCTCTTCTTAAAGAGGAAA	25854
QY	2720	GCAGCTCCTGTAGAAAGCAAGACTTTCTTTTCTGGCTCTTCTTCAACCTTAA	2779
DB	25855	GCAGCTCCTGTAGAAAGCAAGACTTTCTTTTCTGGCTCTTCTTCAACCTTAA	25914
QY	2780	AGCCAGAAACTTGCAGAGTATTTGCTGGGGTGGTTTTATATTTAGGCAATGGGGA	2839
DB	25915	AGCCAGAAACTTGCAGAGTATTTGCTGGGGTGGTTTTATATTTAGGCAATGGGGA	25974
QY	2840	TGGGTGGGAGGGGTTATAGTTTCAAGAGGTTTCTAAGAAATGCTAACAAAGCATT	2899
DB	25975	TGGGTGGGAGGGGTTATAGTTTCAAGAGGTTTCTAAGAAATGCTAACAAAGCATT	26034
QY	2900	TTGGACAATGCTATCCAGCAGGAAAAAAGGATAATATACTGTTTTAAACCTCTTC	2959
DB	26035	TTGGACAATGCTATCCAGCAGGAAAAAAGGATAATATACTGTTTTAAACCTCTTC	26094
QY	2960	TGGGGAATCCAATTATAGTTGCTTTTAAAGCAAGAACAGCAGAGGGTTGTCGC	3019
DB	26095	TGGGGAATCCAATTATAGTTGCTTTTAAAGCAAGAACAGCAGAGGGTTGTCGC	26154
QY	3020	CAGGTAGGATGTCTTAAAGATGGTCCCTTGAATAATGCTTCTGTATCAAGGTA	3079
DB	26155	CAGGTAGGATGTCTTAAAGATGGTCCCTTGAATAATGCTTCTGTATCAAGGTA	26214
QY	3080	CGTATGTGTGCAACAGGCGAAGACTTCTTTTAAATTTCTTCTCTTTTATTAAAC	3139
DB	26215	CGTATGTGTGCAACAGGCGAAGACTTCTTTTAAATTTCTTCTCTTTTATTAAAC	26274
QY	3140	AAATGTGAAGATGGAGGATTTACCTTACAAATCAGACATGGCAACAAATATGGCTTT	3199
DB	26275	AAATGTGAAGATGGAGGATTTACCTTACAAATCAGACATGGCAACAAATATGGCTTT	26334
QY	3200	TGCTTCCATAAAGCAAGTCAATTTTTTAAAGTGTCTCTTACTAAGTCTTTTATTAAAC	3259
DB	26335	TGCTTCCATAAAGCAAGTCAATTTTTTAAAGTGTCTCTTACTAAGTCTTTTATTAAAC	26394
QY	3260	TCTCCTTTATCTATATGAAGTAAAGAGGCGAGTCAATGTAGCAAAATGACGTTAA	3319
DB	26395	TCTCCTTTATCTATATGAAGTAAAGAGGCGAGTCAATGTAGCAAAATGACGTTAA	26454
QY	3320	TATCCCTAGCAGGCGTGTGTTCACTTCCCTGTGATCCCTCTGAGGTATGGCCCATC	3379

Search completed: March 21, 2005, 16:56:22
Job time : 15982 secs

DB	26455	TATCCCTAGCAGAGGCTGTGTTTCACTTCCCTGTGATCCCTCTGAGGTATGGCCCATC	26514
QY	3380	CAAGACTTTTAGGCCATTCTTGTATGGAACACAGATCCCTGCCCTGACTGTCCAGCTATCCT	3439
DB	26515	CAAGACTTTTAGGCCATTCTTGTATGGAACACAGATCCCTGCCCTGACTGTCCAGCTATCCT	36574
QY	3440	GNAAGTGGATCAGATTATAAACTGGATTACATGTAAGTGTGTTGTTGTTCTATCAAC	3499
DB	26575	GNAAGTGGATCAGATTATAAACTGGATTACATGTAAGTGTGTTGTTGTTCTATCAAC	26634
QY	3500	CCCACACAGATTCCCTAAACTTGTCTTCACTTATAGTAAGTGTGTTGTTGTTCTATCAAC	3559
DB	26635	CCCACACAGATTCCCTAAACTTGTCTTCACTTATAGTAAGTGTGTTGTTGTTCTATCAAC	26694
QY	3560	AGCGCATTAAGTCACTTGAAGTATTTGATCCCTTAGATAAGAAATGCAAAATCAGCAGGAAC	3619
DB	26695	AGCGCATTAAGTCACTTGAAGTATTTGATCCCTTAGATAAGAAATGCAAAATCAGCAGGAAC	26754
QY	3620	TGGTCATACAGGTTAAGCACCAGGACCAATAGGATTTTATAGATATATTTAATTTT	3679
DB	26755	TGGTCATACAGGTTAAGCACCAGGACCAATAGGATTTTATAGATATATTTAATTTT	26814
QY	3680	GTTATTTGTTAAGGAGACAAATTTTGGAGAGCAAGCAAAATCTTTTAAAAAATAGTATGAA	3739
DB	26815	GTTATTTGTTAAGGAGACAAATTTTGGAGAGCAAGCAAAATCTTTTAAAAAATAGTATGAA	26874
QY	3740	TGTGAATACTAGAAAAAGATTTAAAAAATAGTATGATGTGTGAGTACTTAGGAAGGAT	3794
DB	26875	TGTGAATACTAGAAAAAGATTTAAAAAATAGTATGATGTGTGAGTACTTAGGAAGGAT	26929

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 21, 2005, 04:05:50 ; Search time 1901 Seconds
(without alignments)
11814.571 Million cell updates/sec

Title: US-10-608-863-1
Perfect score: 3794
Sequence: 1 ataatgacgtccgagaga.....gtgtgagtactaggagagat 3794

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues 8780412

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_16Dec04:*
- 1: Geneseqn1980s:*
 - 2: Geneseqn1990s:*
 - 3: Geneseqn2000s:*
 - 4: Geneseqn2001as:*
 - 5: Geneseqn2001bs:*
 - 6: Geneseqn2002as:*
 - 7: Geneseqn2002bs:*
 - 8: Geneseqn2003as:*
 - 9: Geneseqn2003bs:*
 - 10: Geneseqn2003cs:*
 - 11: Geneseqn2003ds:*
 - 12: Geneseqn2004as:*
 - 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3794	100.0	3794	AD113007	Ad113007 Human NOR
2	3722	98.1	3802	ABK51689	Abk51689 Human nuc
3	3722	98.1	3802	ABK83698	Abk83698 Human cdn
4	3722	98.1	3802	AD157170	Ad157170 Human NOV
5	3679.2	97.0	5828	10 ADE25701	Ad25701 Human cdn
6	3668.2	96.7	5642	8 ABX62937	Abx62937 Human act
7	3653.6	96.3	3762	2 AAT73334	Aat73334 Human neu
8	3352	88.4	4182	10 ADF28866	Adf28866 Human nor
9	2875	75.8	4977	6 ABL61892	Ab161892 Colon ade
10	2875	75.8	4977	6 ABK84229	Abk84229 Human cdn
11	2875	75.8	4977	12 ADL83261	Adl83261 Human PRO
12	2875	75.8	4977	13 ADR14062	Adr14062 Human NF-
13	2689.6	70.9	2714	11 AD131602	Ad131602 Human cdn
14	2500	65.9	4400	2 AAT16151	Aat16151 Apoptotic
15	2500	65.9	4400	6 ABN85437	Abn85437 Murine Ne
16	2500	65.9	4400	10 ADF28867	Adf28867 Rat nor-1
17	2371	62.5	5115	10 ADF28873	Adf28873 Pig nor-1
18	2055	54.2	2055	12 ADP05710	Adp05710 Human nuc
19	1485.6	39.2	1884	6 ABN85436	Abn85436 Murine Ne
20	1022.8	27.0	1806	10 ADF28871	Adf28871 Pig nor-1

21	937.4	24.7	1290	12	ADP05712	Adp05712 Mouse nuc
22	499.4	13.2	3471	12	ADQ22833	Adq22833 Human eof
23	497	13.1	2799	9	ADB84081	Adb84081 Mutant NU
24	497	13.1	3421	9	ADB84043	Adb84043 Human NUR
25	497	13.1	3427	2	AAQ56733	Aaq56733 CDNA sequ
26	497	13.1	3427	4	AAS15573	Aas15573 DNA encod
27	497	13.1	3427	9	ADB84033	Adb84033 Human NUR
28	497	13.1	3427	10	ADH28985	Adh28985 Human chr
29	497	13.1	3427	10	ACC58210	Acc58210 Human NUR
30	497	13.1	3427	11	AD132074	Ad132074 Human cdn
31	497	13.1	3427	12	ADH75169	Adh75169 Human TIN
32	497	13.1	3427	13	ADP23739	Adp23739 PRO polyP
33	497	13.1	3436	10	ADE25697	Ad25697 Human cdn
34	496	13.1	2154	9	ADB84053	Adb84053 Rat NURR1
35	495.4	13.1	1797	4	AAC67273	Aac67273 Human nuc
36	495.4	13.1	1797	12	ADP05706	Adp05706 Human nuc
37	492.8	13.0	2002	13	ADR14978	Adr14978 Rat elect
38	492.4	13.0	2247	9	ADB84058	Adb84058 Murine NU
39	492.4	13.0	2458	9	ADB84032	Adb84032 Murine NU
40	490.8	12.9	1797	12	ADP05708	Adp05708 Mouse nuc
41	486.4	12.8	1797	3	AAA95361	Aaa95361 Rat Nurr1
42	486	12.8	1821	5	AAS14641	Aas14641 Mouse cdn
43	483.2	12.7	2559	10	ACC58209	Acc58209 Rat Nurr1
44	469.4	12.4	2469	10	ADF28894	Adf28894 Human nur
45	469.4	12.4	2469	10	ADK61137	Adk61137 Ovarian c

ALIGNMENTS

RESULT 1
AD113007
ID AD113007 standard; cdna; 3794 BP.
XX
AC AD113007;
XX
DT 22-APR-2004 (first entry)
XX
DE Human NOR-1 (MINOR) cdna sequence SeqID 1.

KW human; gene; ss; allergic disease; NOR-1; MINOR; eosinophil;
KW atopic dermatitis; antiallergic; antiinflammatory; dermatological.

OS Homo sapiens.

XX Key Location/Qualifiers
FT CDS 730..2610
FT /tag= a
FT /product= "NOR-1 protein"

XX WO2004003198-A1.

XX 08-JAN-2004.

XX 27-JUN-2003; 2003WO-JP008199.

XX 27-JUN-2002; 2002JP-00188490.

XX (GENO-) GENOX RES INC.

XX (NIGE-) JAPAN GEN AGENCY NATION.

XX Hashida R, Kagaya S, Yayoi Y, Sugita Y, Saito H;

XX WPI; 2004-083057/08.

XX P-PSDB; AD113008.

XX Examining allergic diseases e.g. atopic dermatitis by differential display based on gene expression of NOR-1 receptor protein, also applicable in screening compounds for treatment of allergic diseases.

XX Example 1; SEQ ID NO 1; 155pp; Japanese.

XX This invention relates to a novel method for examining allergic diseases

that comprises comparing the expression levels of a gene encoding the NOR-1 receptor protein between patients and healthy individuals. Specifically, the NOR-1 gene, also referred to as MINOR, is expressed in the specialist white blood cells known as eosinophils and is involved in mediating an allergic reaction. The present invention describes a differential display method that can identify the expression level of this gene in order to identify its usefulness in diagnosing allergic diseases such as atopic dermatitis. Furthermore, compositions can also be used to screen compounds for the treatment of allergic diseases. Accordingly, they exhibit various activities including antiallergic, anti-inflammatory and dermatological. This polynucleotide sequence is the human NOR-1 cDNA sequence of the invention.

SQ	Sequence	3794 BP; 917 A; 1140 C; 910 G; 827 T; 0 U; 0 Other;
	Query Match	100.0%; Score 3794; DB 12; Length 3794;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 3794; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	ATAAATGACGTCCGAGAGCGAGCGGAACCGCAGCCGGGAGAGCGGAGTCCTCTCGCT
Db	1	ATAAATGACGTCCGAGAGCGAGCGGAACCGCAGCCGGGAGAGCGGAGTCCTCTCGCT
Qy	61	CCGCGCCCCACCCCTCCAGCTCTCGCTCTCTCCGCTCCCTCCGCTCCCATACACAGACGCGCTCA
Db	61	CCGCGCCCCACCCCTCCAGCTCTCTCGCTCTCTCCGCTCCCATACACAGACGCGCTCA
Qy	121	CACCGCTCCTTCATCTGGCACACACAGACACAAGCGGCGCACACAGGTCCTCGCACACAC
Db	121	CACCGCTCCTTCATCTGGCACACACAGACACAAGCGGCGCACACAGGTCCTCGCACACAC
Qy	181	TTTGCCTCTCCCGGGCGCTACACCCCTCTTGCCCTTGAGCCCTTGCCTGAGCGGCGGCG
Db	181	TTTGCCTCTCCCGGGCGCTACACCCCTCTTGCCCTTGAGCCCTTGCCTGAGCGGCGGCG
Qy	241	CCGACGCTGGAGCGCCCTCCCGGGCTCACTTTGCAAAGCTGACGGTCCCGGCAGTGCGCG
Db	241	CCGACGCTGGAGCGCCCTCCCGGGCTCACTTTGCAAAGCTGACGGTCCCGGCAGTGCGCG
Qy	301	TGAGGTGGGAACAGGGCGGCGATCTCTCCCTCTGGTCAACGCCAAGACGACGACGCGCG
Db	301	TGAGGTGGGAACAGGGCGGCGATCTCTCCCTCTGGTCAACGCCAAGACGACGACGCGCG
Qy	361	CGGAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATCGCAGACTCCCCACACGACGCT
Db	361	CGGAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATCGCAGACTCCCCACACGACGCT
Qy	421	CACCGCTCTCGGAGCGCGTGGCTTGTACACCGCAGCCCTTCCGGACAGCAGCTGTGA
Db	421	CACCGCTCTCGGAGCGCGTGGCTTGTACACCGCAGCCCTTCCGGACAGCAGCTGTGA
Qy	481	CTCCCCCAGTGCAGATTTCGGGACAGCTCTCTAGAAACTCGCTCTAAAGACGGACCG
Db	481	CTCCCCCAGTGCAGATTTCGGGACAGCTCTCTAGAAACTCGCTCTAAAGACGGACCG
Qy	541	CCACAGACTCAAAGGCCACTTCGGAAGAGGCGACCGCGAACCGCGGCCCTTGAGCCT
Db	541	CCACAGACTCAAAGGCCACTTCGGAAGAGGCGACCGCGAACCGCGGCCCTTGAGCCT
Qy	601	GGACCTTACGCGTGCGGGACGCACTCCGCGCGCTTCGCTTCGCGGACGCTCGCTCCT
Db	601	GGACCTTACGCGTGCGGGACGCACTCCGCGCGCTTCGCTTCGCGGACGCTCGCTCCT
Qy	661	CCTACACTCTACGCTTCGGAGAGACCCCGACCCCACTTCAGCGGCGCAAGATAC
Db	661	CCTACACTCTACGCTTCGGAGAGACCCCGACCCCACTTCAGCGGCGCAAGATAC
Qy	721	CCTCCAGATATGCCCTGCGTCCAAAGCCAAATATAGCCCTTCCCCTCCAGGTTCCAGTTAT
Db	721	CCTCCAGATATGCCCTGCGTCCAAAGCCAAATATAGCCCTTCCCCTCCAGGTTCCAGTTAT
Qy	781	GCGGCGGACACATACAGCTCGGAATACACCAACGGAGATCATGAACCCCGACTACAACAG
	840	

XX WO200187923-A1.
 XX 22-NOV-2001.
 XX 11-MAY-2001; 2001WO-US015311.
 XX 12-MAY-2000; 2000US-0203645P.
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 XX Murphy E, Conneely OM, Fitzgerald O, Bresihan B;
 XX WPI; 2002-075311/10.
 XX Treating inflammatory immune disease such as arthritis, comprises
 PT suppressing expression level of NURR subfamily of nuclear transcription
 PT factors, or corticotropin releasing hormone receptor.
 XX
 XX Claim 10; Page; 123pp; English.
 XX The present invention relates to a new method of treating an organism for
 CC an inflammatory immune disease. The method of the invention comprises
 CC reducing expression of a NURR subfamily nucleic acid sequence or
 CC corticotropin releasing hormone (CRH) receptor nucleic acid sequence,
 CC inhibiting transcriptional activity of a NURR superfamily member/CRH
 CC receptor amino acid sequence, or reducing the level of NURR superfamily
 CC member/CRH receptor sequence. The method is useful for treating an
 CC organism for an inflammatory immune diseases such as chronic inflammatory
 CC joint disease, preferably arthritis, selected from rheumatoid arthritis,
 CC psoriatic arthritis and sarcoid arthritis, ulcerative colitis and
 CC thyroiditis. The method is also useful for screening a compound that
 CC interferes with interaction of a NURR subfamily polypeptide with a
 CC ligand, or identifying a compound for the treatment of an inflammatory
 CC immune response. The agonist of the invention is useful for inhibiting
 CC transcriptional activity of nuclear receptor polypeptide and the
 CC antagonist is useful for decreasing the expression of a NURR subfamily
 CC member. The present nucleic acid sequence represents the human nuclear
 CC receptor NOR1 gene of the invention. Note: The present sequence is not
 CC shown in the specification but was obtained from Genbank (1651190)
 XX
 XX Sequence 3802 BP; 919 A; 1141 C; 912 G; 830 T; 0 U; 0 Other;
 SQ
 Query Match 98.1%; Score 3722; DB 6; Length 3802;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;
 QY 1 ATAAATGAGTCCGAGAGAGCGAGACGCGAGCGGAGCGGAGTCTCTGCT 60
 DB |||||
 QY 1 ATAAATGAGTCCGAGAGAGCGAGACGCGAGCGGAGCGGAGTCTCTGCT 60
 DB |||||
 QY 61 CCGGCCCCCACCCTCCAGTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
 DB |||||
 QY 61 CCGGCCCCCACCCTCCAGTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
 DB |||||
 QY 121 CACCGCTCCCTCACTCGACACACAGACACAAGCGCGCACACAGCTCCG--CACAC 178
 DB |||||
 QY 121 CACCGCTCCCTCACTCGAACAACACAGACACAAGCGCGCACACAGGCTCCGACAC 180
 DB |||||
 QY 179 ACTTCGCTCTCCGCGGCTCACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 238
 DB |||||
 QY 181 ACTTCGCTCTCCGCGGCTCACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 DB |||||
 QY 239 CGCGGAGTGGACGCGCTCTCCGCGCTCACTTTGCAACGCTGACGCTGCGCGAGTGGC 298
 DB |||||
 QY 241 CGCGGAGTGGACGCGCTCTCCGCGCTCACTTTGCAACGCTGACGCTGCGCGAGTGGC 300
 DB |||||
 QY 299 CGTGGAGTGGAAACAGCGCGGCTATCTCTCCCTGCTGCTACAGCCCAAGCCAGGACGCC 358
 DB |||||
 QY 301 CGTGGAGTGGAAACAGCGCGGCTATCTCTCCCTGCTGCTACAGCCCAAGCCAGGACGCC 360
 DB |||||
 QY 359 CGCGGAGTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 418
 DB |||||

Db 361 CGCGAACCCTCTCGGCTGTGCTCTCTCCCATGATCGGATCGGAGATCCCCCACAGCGG 420
 QY 419 CTCACCGCTCCGGAGCGCTGGGCTTGTACACCGCAGCGCTTCCGGGACAGCAGCTGT 478
 DB |||||
 DB 421 CTCACCGCTCCGGAGCGCTGGGCTTGTACACCGCAGCGCTTCCGGGACAGCAGCTGT 480
 QY 479 GACTCCCCCAGTGCAGATTTTGGGACAGCTCTCTAGAACTGCTCTTAAGACGGAAC 538
 DB |||||
 DB 481 GACTCCCCCAGTGCAGATTTTGGGACAGCTCTCTAGAACTGCTCTTAAGACGGAAC 540
 QY 539 CGCACAGCACTCAAAAGCCCACTCGGAAGAGGCGAGCCCGGCAAGCCCGGCGCTGAGC 598
 DB |||||
 DB 541 CGCACAGCACTCAAAAGCCCACTCGGAAGAGGCGAGCCCGGCAAGCCCGGCGCTGAGC 600
 QY 599 CTGACCCCTTAGCGGTGCGGGGAGCACTGCGGGGCTTTCGCTCGCGGAGCTCGCTC 658
 DB |||||
 DB 601 CTGACCCCTTAGCGGTGCGGGGAGCACTGCGGGGCTTTCGCTCGCGGAGCTCGCTC 660
 QY 659 CTCTTACACTCTCAGCTTCGCTGAGAGAGCCCGCAGCCCACTTTCAGCGCGCAAGAT 718
 DB |||||
 DB 661 CTCTTACACTCTCAGCTTCGCTGAGAGAGCCCGCAGCCCACTTTCAGCGCGCAAGAT 720
 QY 719 ACCCTCCAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778
 DB |||||
 DB 721 ACCCTCCAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 779 ATGCGGCGCAGACATACAGCTCGGAATACACACGAGATCATGAACCCCGACTACCA 838
 DB |||||
 DB 781 ATGCGGCGCAGACATACAGCTCGGAATACACACGAGATCATGAACCCCGACTACCA 840
 QY 839 AGTGACCATGACCTTGGGAGCTAGATACAGGCTACAGCCACAGCTCCCTGCCCA 898
 DB |||||
 DB 841 AGTGACCATGACCTTGGGAGCTAGATACAGGCTACAGCCACAGCTCCCTGCCCA 900
 QY 899 GCATCAGTACCTTCTGAGAGGCTACTCGAGCACTACGAACTCAAGCTTCTCTCGGT 958
 DB |||||
 DB 901 GCATCAGTACCTTCTGAGAGGCTACTCGAGCACTACGAACTCAAGCTTCTCTCGGT 960
 QY 959 ACCAAATGACGCGCTTGTATCAAGTGGAGGCGGCGCGCCAGCTACCATACC 1018
 DB |||||
 DB 961 ACCAAATGACGCGCTTGTATCAAGTGGAGGCGGCGCGCCAGCTACCATACC 1020
 QY 1019 ATCACCAC 1078
 DB |||||
 DB 1021 ATCACCAC 1080
 QY 1079 TTCTTCCAGCTTCCAGCCCGGAGGAGTGTCTCCCGCAGCACCTTCTCTCAAGC 1138
 DB |||||
 DB 1081 TTCTTCCAGCTTCCAGCCCGGAGGAGTGTCTCCCGCAGCACCTTCTCTCAAGC 1140
 QY 1139 AGTCCCAACCTTCCAGCCCGGAGGAGTGTCTCCCGCAGCACCTTCTCTCAAGC 1198
 DB |||||
 DB 1141 AGTCCCAACCTTCCAGCCCGGAGGAGTGTCTCCCGCAGCACCTTCTCTCAAGC 1200
 QY 1199 AGGAGGCTGCTGCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1258
 DB |||||
 DB 1201 AGGAGGCTGCTGCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 QY 1259 TGAAGCGCTTCCCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1318
 DB |||||
 DB 1261 TGAAGCGCTTCCCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 QY 1319 CGCCGCATCCCG 1378
 DB |||||
 DB 1321 CGCCGCATCCCG 1380
 QY 1379 CGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1438
 DB |||||
 DB 1381 CGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 QY 1439 CGCTTTGAGAGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1498
 DB |||||
 DB 1441 CGCTTTGAGAGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500

1499 CGCCTCTCGGCTCAGCCCTCCCTACCGCGTCCAGCCTGCTGGCGAGAGTCCAGCC 1558
1501 CGCCTCTCGGCTCAGCCCTCCCTACCGCGTCCAGCCTGCTGGCGAGAGTCCAGCC 1560
1559 TGCCTGCGCGCCAGCAGAGTCTGTCGTCTGGCGAGGCGACGTGTGCGGTGCGGGG 1618
1561 TGCCTGCGCGCCAGCAGAGTCTGTCGTCTGGCGAGGCGACGTGTGCGGTGCGGGG 1620
1619 ACAACGCGCTGCCAGCACTACGGCGTGCAGCTGCGAGGGCTGCAAGGGCTTTTCA 1678
1621 ACAACGCGCTGCCAGCACTACGGCGTGCAGCTGCGAGGGCTGCAAGGGCTTTTCA 1680
1679 AGAAGACAGTGCAGAAAATGCAAAATATGTTGCTGCGCAATTAATAACTCCCACTAG 1738
1681 AGAAGACAGTGCAGAAAATGCAAAATATGTTGCTGCGCAATTAATAACTCCCACTAG 1740
1739 ACAAGACAGTGCAGAAACCGATGTCAGTACTGTCGATTTTCAAGTGTCTCAGTGTGGAA 1798
1741 ACAAGACAGTGCAGAAACCGATGTCAGTACTGTCGATTTTCAAGTGTCTCAGTGTGGAA 1800
1799 TGGTAAAAAGATGTCGTGTCAGATAGTCTGAAAGGGAGGAGTGTCTGCTCCCTTCCA 1858
1801 TGGTAAAAAGATGTCGTGTCAGATAGTCTGAAAGGGAGGAGTGTCTGCTCCCTTCCA 1860
1859 AACCAAGAGCCATTTACAAAGGAACCTTCTCAGCCCTCTCCAATCTCCTCCCAATCT 1918
1861 AACCAAGAGCCATTTACAAAGGAACCTTCTCAGCCCTCTCCAATCTCCTCCCAATCT 1920
1919 GCATGATGAATGCCCTTGTCCGAGCTTTAAACAGCTCAACACCCAGAGATCTTGATTT 1978
1921 GCATGATGAATGCCCTTGTCCGAGCTTTAAACAGCTCAACACCCAGAGATCTTGATTT 1980
1979 CCAGATAGTCTCCCACTGACAGGCTGTCGAGGCAAGATGTCGAGCAATGTCGAACAT 2038
1981 CCAGATAGTCTCCCACTGACAGGCTGTCGAGGCAAGATGTCGAGCAATGTCGAACAT 2040
2039 TCTACAACTCTGACAGCTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 2098
2041 TCTACAACTCTGACAGCTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
2099 GATTTACTGATCTCCCAAGAGATCAGACATTAATTAATCAATCAGCTTTTGGAGC 2158
2101 GATTTACTGATCTCCCAAGAGATCAGACATTAATTAATCAATCAGCTTTTGGAGC 2160
2159 TGTGTGCTCAGAGCTTCCATCAGGTCAGGCTGTCGAGGCAAGATGATGATGATGATGAT 2218
2161 TGTGTGCTCAGAGCTTCCATCAGGTCAGGCTGTCGAGGCAAGATGATGATGATGATGAT 2220
2219 ATGGACTGTCTGCAATCAGTCTGAGTGCCTTGGTGGATTTGGGAGTGGCTCGACTCTA 2278
2221 ATGGACTGTCTGCAATCAGTCTGAGTGCCTTGGTGGATTTGGGAGTGGCTCGACTCTA 2280
2279 TTAAGACTTTTCCCTTAAATTTGAGAGCTGAACTTGAATCCAGGCTTAGGCTGCCC 2338
2281 TTAAGACTTTTCCCTTAAATTTGAGAGCTGAACTTGAATCCAGGCTTAGGCTGCCC 2340
2339 TGTGAGCACTGAGCATGATCAGAGAAAGACATGGGTAAAGAAACCAAGAGAGTCCGAG 2398
2341 TGTGAGCACTGAGCATGATCAGAGAAAGACATGGGTAAAGAAACCAAGAGAGTCCGAG 2400
2399 AGCTATGCAACAGATCAGAGAGTAAAGAGCCAGAGTAAAGGAGAGGCTGCTGG 2458
2401 AGCTATGCAACAGATCAGAGAGTAAAGAGCCAGAGTAAAGGAGAGGCTGCTGG 2460
2459 AGCCCAAGAGTCCAGGCTGCTGGTCCCTGCTAGAGTGAAGATCTGACCCCTGG 2518
2461 AACCAAGAGTCCAGGCTGCTGGTCCCTGCTAGAGTGAAGATCTGACCCCTGG 2520
2519 GCCTCAGGCACTCTTACCTGAGCTGGAAGCTTGGTGTCTCCACCTTCCATCATTTG 2578
2521 GCCTCAGGCACTCTTACCTGAGCTGGAAGCTTGGTGTCTCCACCTTCCATCATTTG 2580

2579 ACAAGCTCTTCTCTGACACCCCTACCTTTCTAATCAGAGGAGTGGAGAGTGCCT 2638
2581 ACAAGCTCTTCTCTGACACCCCTACCTTTCTAATCAGAGGAGTGGAGAGTGCCT 2640
2639 CCTCTCTAGCACCTGCTTGTACGACCAAGAGGATAGGTTTGAACCTATCATTTCC 2698
2641 CCTCTCTAGCACCTGCTTGTACGACCAAGAGGATAGGTTTGAACCTATCATTTCC 2700
2699 TGTCTCTCTTAAAGAGGAAAAAGCAGCTCTGTAGAAAGCAAGACCTTTCTTTTCTG 2758
2701 TGTCTCTCTTAAAGAGGAAAAAGCAGCTCTGTAGAAAGCAAGACCTTTCTTTTCTG 2760
2759 GCTCTTTTCTTACAACTTAAAGCCAGAAACTTCCAGAGTATTTGTTGGGTTGTT 2818
2761 GCTCTTTTCTTACAACTTAAAGCCAGAAACTTCCAGAGTATTTGTTGGGTTGTT 2820
2819 TTATATTTAGGCATTTGGGGATGGGGTGGGGGTTATAGTTTATCAGGGTCTTCTAA 2878
2821 TTATATTTAGGCATTTGGGGATGGGGTGGGGGTTATAGTTTATCAGGGTCTTCTAA 2880
2879 GAAATGCTAAACAAAGCACTTTTGAACAATGCTATCCAGCAGGAAAAAAGGATAATA 2938
2881 GAAATGCTAAACAAAGCACTTTTGAACAATGCTATCCAGCAGGAAAAAAGGATAATA 2940
2939 TAACTGTTTAAACCTCTTCTGGGAATCCAAATTTATAGTTGCTTTGTTTAAACAA 2998
2941 TAACTGTTTAAACCTCTTCTGGGAATCCAAATTTATAGTTGCTTTGTTTAAACAA 3000
2999 GAAACAGCAAGGTTGTTGCGCAGGTTAGGATGCTCTTAAAGATTTGGTCCCTTGAAT 3058
3001 GAAACAGCAAGGTTGTTGCGCAGGTTAGGATGCTCTTAAAGATTTGGTCCCTTGAAT 3060
3059 ATGCTTCTGATCAAGGTAAGTATGTTGTCGCAACAAAGCAGGAAACTTCTTTTAAAT 3118
3061 ATGCTTCTGATCAAGGTAAGTATGTTGTCGCAACAAAGCAGGAAACTTCTTTTAAAT 3120
3119 TCCTTCTTCTTATTTTAAACAAATGTTGTAAGATGAGGATTAACCTACAAATCAGACAT 3178
3121 TCCTTCTTCTTATTTTAAACAAATGTTGTAAGATGAGGATTAACCTACAAATCAGACAT 3180
3179 GCAAAAAAATAAGCTGTTTGTCTTCAATAAACAAGTGCATTTTAAAGTGTCT 3238
3181 GCAAAAAAATAAGCTGTTTGTCTTCAATAAACAAGTGCATTTTAAAGTGTCT 3240
3239 TACTAAGTCTTGTATTAACCTCTCTTTATTTGTAATAAAGGAGGAGTCA 3298
3241 TACTAAGTCTTGTATTAACCTCTCTTTATTTGTAATAAAGGAGGAGTCA 3300
3299 TGTAGCAATGACAGCTTAATATCCCTAGCAGAGGCTGTGTACCTTCCCTGTCGATC 3358
3301 TGTAGCAATGACAGCTTAATATCCCTAGCAGAGGCTGTGTACCTTCCCTGTCGATC 3360
3359 CCTTCTGAGGTATGGCCCATCCAGACTTTTAGGCAATTTTGGTGAACAGATCCCTG 3418
3361 CCTTCTGAGGTATGGCCCATCCAGACTTTTAGGCAATTTTGGTGAACAGATCCCTG 3420
3419 CCTGACTGTCCAGTATCCTGAAAGTGGATCAGATTAATACTGGATTAACGATG 3478
3421 CCTGACTGTCCAGTATCCTGAAAGTGGATCAGATTAATACTGGATTAACGATG 3480
3479 TTTTGGTGTGTTCTATCAACCCAGCAGAGTTCCTAACTGCTTCAGTTATAGTAAC 3538
3481 TTTTGGTGTGTTCTATCAACCCAGCAGAGTTCCTAACTGCTTCAGTTATAGTAAC 3540
3539 TGACTGTGATATTTCAATCAGAGGCCATAGCTCAGTTGAGTATTTGATCCCTAGATAAG 3598
3541 TGACTGTGATATTTCAATCAGAGGCCATAGCTCAGTTGAGTATTTGATCCCTAGATAAG 3600
3599 AACATGCAAAATCAGCAGGAACTGTCATACAGGGTGAAGCAGGAGCAATTAAGGATTT 3658
3601 AACATGCAAAATCAGCAGGAACTGTCATACAGGGTGAAGCAGGAGCAATTAAGGATTT 3660
3659 TATAGATATAATTTAAATTTTGTGTA - TTGGTTAAGGAGA - CAATTTTGGAGAGCAGCA 3715

Db 3661 TATAGATATAATTTTAAATTTTGGTAAATGGGTAAAGGAGACCAATTTTGGAGAGCA 3720
 QY 3716 AA---TCTTTTAAAAATAGTATGAACTGTAATGTAAGTAAAGATTTTAAAAAATAGTAT 3772
 Db 3721 AATCTTCTTTTAAAAAATAGTATGAACTGTAATGTAAGTAAAGATTTTAAAAAATAGTAT 3780
 QY 3773 GAGTGTGAGTACTAGGAAGAT 3794
 Db 3781 GAGTGTGAGTACTAGGAAGAT 3802

RESULT 3

ABK83698

ID ABK83698 standard; cDNA; 3802 BP.

XX AC ABK83698;

XX DT 14-AUG-2002 (first entry)

XX DE Human cDNA differentially expressed in granulocytic cells #269.

XX KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.

XX OS Homo sapiens.

XX PN WO200228999-A2.

XX PD 11-APR-2002.

XX PF 03-OCT-2001; 2001WO-US030821.

XX PR 03-OCT-2000; 2000US-0237189P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;

XX DR WPI; 2002-435328/46.

XX PT Detecting granulocyte activation by detecting differential expression of
 PT genes associated with granulocyte activation, which serves as diagnostic
 PT markers that is useful for monitoring disease states and drug toxicity.

XX PS Claim 1; SEQ ID NO 269; 114pp; English.

XX CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing the
 CC expression level to an expression level in an unactivated GC, where
 CC differential expression of Gs is indicative of GCA. Also included are
 CC modulating (M2) GCA by contacting GC with an agent that alters the
 CC expression of at least one gene in Gs; (2) screening (M3) for an agent
 CC capable of modulating GCA or an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease using the gene expression
 CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease, by detecting the level of
 CC expression in a sample of the tissue of gene(s) from Gs, where the level
 CC of expression of the gene is indicative of inflammation; (4) treating
 CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease, by contacting a tissue having inflammation with an
 CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
 CC is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful

CC for screening an agent capable of modulating GCA preferably in an
 CC inflammation in a tissue; M4 is useful for detecting an inflammation
 CC (especially chronic) in a tissue, an allergic response in a subject,
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, periodontal disease; also bacterial
 CC infection, viral infection, parasitic infection, protozoal infection,
 CC fungal infection and M5 is useful for treating one of the above
 CC conditions. The present sequence represents a gene differentially
 CC expressed in granulocytes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 3802 BP, 919 A; 1141 C; 912 G; 830 T; 0 U; 0 Other;

Query Match 98.1%; Score 3722; DB 6; Length 3802;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY 1 ATAAATGACGTGCGGAGAGCGGAGCGGAAACGCGAGCGCGGAGCGGAGTCTCTGCT 60
 Db 1 ATAAATGACGTGCGGAGAGCGGAGCGGAAACGCGAGCGCGGAGCGGAGTCTCTGCT 60
 QY 61 CCGCGCCCCACCCCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 120
 Db 61 CCGCGCCCCACCCCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 120
 QY 121 CACCGGCT 178
 Db 121 CACCGGCT 180
 QY 179 ACTTCGCTCTCCCGCGGCTCACACCCCTCTTCGCTCAGCGCTTTCGCGGTGCGCGGG 238
 Db 181 ACTTCGCTCTCCCGCGGCTCACACCCCTCTTCGCTCAGCGCTTTCGCGGTGCGCGGG 240
 QY 239 CGCGCAGCTGGAGCGCCCTCCCGGGCTCACTTTGCAACGCTGACGCTGCGCGGAGTGGC 298
 Db 241 CGCGCAGCTGGAGCGCCCTCCCGGGCTCACTTTGCAACGCTGACGCTGCGCGGAGTGGC 300
 QY 299 CGTGAGGTGGGAAACAGCGCGGCGCATCTCTCCCTCTGTGTACAGCCCAAGCCAGGACGCC 358
 Db 301 CGTGAGGTGGGAAACAGCGCGGCGCATCTCTCCCTCTGTGTACAGCCCAAGCCAGGACGCC 360
 QY 359 CGCGGAACTCTCGGCTGTCTCTCCATGAGTCGGGATCGCAGCATCCCCCAGCAGCG 418
 Db 361 CGCGGAACTCTCGGCTGTCTCTCCATGAGTCGGGATCGCAGCATCCCCCAGCAGCG 420
 QY 419 CTCACCGGCTCCGGGAGCGGCTGGGCTTGTATACCGCAGCCCTTCGCGGACAGCAGCTGT 478
 Db 421 CTCACCGGCTCCGGGAGCGGCTGGGCTTGTATACCGCAGCCCTTCGCGGACAGCAGCTGT 480
 QY 479 GACTCCCCCAGGTGCAGATTTGGGAGCAGCTCTCTAGAACTCGCTCTAAAGAGCGGAAC 538
 Db 481 GACTCCCCCAGGTGCAGATTTGGGAGCAGCTCTCTAGAACTCGCTCTAAAGAGCGGAAC 540
 QY 539 CGCCACAGCACTCAAAAGCCCACTCGGAGAGGGGAGCGCCGCGCAAGCCCGGCCCTGAGC 598
 Db 541 CGCCACAGCACTCAAAAGCCCACTCGGAGAGGGGAGCGCCGCGCAAGCCCGGCCCTGAGC 600
 QY 599 CTGAGCCCTTAGCGGTGCGGGGAGCAGCTGCGGGGCTTTCGCTCTCCCGGACGTCCGCTC 658
 Db 601 CTGAGCCCTTAGCGGTGCGGGGAGCAGCTGCGGGGCTTTCGCTCTCCCGGACGTCCGCTC 660
 QY 659 CTCCTACACTCTACGCTTCGCTGAGAGACCCCGAGCCCACTTTCAGCGCGCAAGAT 718
 Db 661 CTCCTACACTCTACGCTTCGCTGAGAGACCCCGAGCCCACTTTCAGCGCGCAAGAT 720
 QY 719 ACCCTCCAGATATGCGCTCGGTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCAGT 778
 Db 721 ACCCTCCAGATATGCGCTCGGTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCAGT 780

QY 779 ATGCGGCGAGATACAGCTCGGAATACACACGAGATCATGAACCCGACTACACCA 838
DB |||||
781 ATGCGGCGAGATACAGCTCGGAATACACACGAGATCATGAACCCGACTACACCA 840
QY |||||
839 AGCTGACCATGGAACCTTGGCAGCACTGAGATACGGCTACAGCCACACAGCTCCCTGCCCA 898
DB |||||
841 AGCTGACCATGGAACCTTGGCAGCACTGAGATACGGCTACAGCCACACAGCTCCCTGCCCA 900
QY |||||
899 GCATCAGTACCTTTCGAGGGGCTACTCGAGCAAACTACGAACTCAAGCCCTTCCTGCGTGT 958
DB |||||
901 GCATCAGTACCTTTCGAGGGGCTACTCGAGCAAACTACGAACTCAAGCCCTTCCTGCGTGT 960
QY |||||
959 ACCAAATGACGGCCCTTGATCAAAAGTGGAGAGGGGGCGGGCCAGCTACCATCACCC 1018
DB |||||
961 ACCAAATGACGGCCCTTGATCAAAAGTGGAGAGGGGGCGGGCCAGCTACCATCACCC 1020
QY |||||
1019 ATCACCACCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1078
DB |||||
1021 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1080
QY |||||
1079 TTCTCTCAGCCTCCAGCCTCGGAGGACGAGGTGCTGCCAGCACTCCATGTACTTCAAGC 1138
DB |||||
1081 TTCTCTCAGCCTCCAGCCTCGGAGGACGAGGTGCTGCCAGCACTCCATGTACTTCAAGC 1140
QY |||||
1139 AGTCCCCACGGTCCACCCCAACCAACCGCGGCTTCCCCCGGAGGGGGGGGTTATGGG 1198
DB |||||
1141 AGTCCCCACGGTCCACCCCAACCAACCGCGGCTTCCCCCGGAGGGGGGGGTTATGGG 1200
QY |||||
1199 ACAGGACACTGCCCCGCGCGCGCTGCATCGACCCCGCGCGCTGCTGGACCCGCGCA 1258
DB |||||
1201 ACAGGACACTGCCCCGCGCGCGCTGCATCGACCCCGCGCGCTGCTGGACCCGCGCA 1260
QY |||||
1259 TGAAGCGGTGCCACCGGTGGCGCGCGCTTCCGCTCTTCACTTCAAGCCCTCGC 1318
DB |||||
1261 TGAAGCGGTGCCACCGGTGGCGCGCGCTTCCGCTCTTCACTTCAAGCCCTCGC 1320
QY |||||
1319 CCGCGCATCCCCCG 1378
DB |||||
1321 CCGCGCATCCCCCG 1380
QY |||||
1379 CCGCTGCGCGCTCAGCTGCGCTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1438
DB |||||
1381 CCGCTGCGCGCTCAGCTGCGCTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY |||||
1439 CCGCTGAGAGCCACCGGTACGGCTGCGCGCTGCGCAAGAGGGCGCGCGCGCGCGCGCGCGCG 1498
DB |||||
1441 CCGTTGAGGGCCACCGGTACGGCTGCGCGCTGGCCAGAGGGGGCGCGCGCGCGCGCGCGCG 1500
QY |||||
1499 CCGCTCTCGGCTCAGCGCTCCCTACCGCTCGAGCTGCTGGCGGAGAGTCCAGGCC 1558
DB |||||
1501 CCGCTCTCGGCTCAGCGCTCCCTACCGCTCGAGCTGCTGGCGGAGAGTCCAGGCC 1560
QY |||||
1559 TCGCGTCCGCGCGCGAGGAGCTGCTGCTGGCGAGGGGACGCTGCGCGGTGCGGGG 1618
DB |||||
1561 TCGCGTCCGCGCGCGAGGAGCTGCTGCTGGCGAGGGGACGCTGCTGCGGTGCGGGG 1620
QY |||||
1619 ACAACGCGCGCTCCAGCACTACGGCTGCGAACCTGCGAGGGCTGCAAGGGCTTTTCA 1678
DB |||||
1621 ACAACGCGCGCTCCAGCACTACGGCTGCGAACCTGCGAGGGCTGCAAGGGCTTTTCA 1680
QY |||||
1679 AGAGAACAGTGAGAAAAATGCAAAATATGTTTGCCTGCGCAATAAAAACTGCCAGTAG 1738
DB |||||
1681 AGAGAACAGTGAGAAAAATGCAAAATATGTTTGCCTGCGCAATAAAAACTGCCAGTAG 1740
QY |||||
1739 ACAAGAGAGTGGAAACCGATGCTAGTACTGCTGATTTTCAAGAGTGTCTCAGTGTGGAA 1798
DB |||||
1741 ACAAGAGAGTGGAAACCGATGCTAGTACTGCTGATTTTCAAGAGTGTCTCAGTGTGGAA 1800
QY |||||
1799 TGGTAAAGAAAGTTGTCCGTACAGATAGTCTGAAGAGGAGGAGAGTGTGCTGCGCTTCCA 1858
DB |||||
1801 TGGTAAAGAAAGTTGTCCGTACAGATAGTCTGAAGAGGAGGAGAGTGTGCTGCGCTTCCA 1860

QY 1859 AACCAAGAGCCATTACAAAGGACCTTCTCAGCCCTCTCCACCTTCTCTCCCAATCT 1918
DB |||||
1861 AACCAAGAGCCATTACAAAGGACCTTCTCAGCCCTCTCCACCTTCTCTCCCAATCT 1920
QY |||||
1919 GCATGATGAATGAGCCCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTAAT 1978
DB |||||
1921 GCATGATGAATGAGCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTAAT 1980
QY |||||
1979 CCAGATATGTCCTCCACTGACACAGGCTGTGACAGGACAGATGCTGAGCATGTGCAACAAT 2038
DB |||||
1981 CCAGATATGTCCTCCACTGACACAGGCTGTGACAGGACAGATGCTGAGCATGTGCAACAAT 2040
QY |||||
2039 TCTCAACCTCTGACAGCCTCCATTTGATGATCCAGAGCTGGGAGAGAAAGATTTCCGG 2098
DB |||||
2041 TCTCAACCTCTGACAGCCTCCATTTGATGATCCAGAGCTGGGAGAGAAAGATTTCCGG 2100
QY |||||
2099 GATTTACTGATCTCCCCAAAGAGATCAGACATTAATTGAATCAGCCTTTTTCGAGC 2158
DB |||||
2101 GATTTACTGATCTCCCCAAAGAGATCAGACATTAATTGAATCAGCCTTTTTCGAGC 2160
QY |||||
2159 TGTGTGCTCAGACTTTCCATCAGGTCAAAACACTGCTGAAGATTAAGTTTGTGTTCTGCA 2218
DB |||||
2161 TGTGTGCTCAGACTTTCCATCAGGTCAAAACACTGCTGAAGATTAAGTTTGTGTTCTGCA 2220
QY |||||
2219 ATGGATGTCCTGCACTCGACTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTA 2278
DB |||||
2221 ATGGATGTCCTGCACTCGACTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTA 2280
QY |||||
2279 TTAAGAGCTTTTCTTAAATTTGAGAGCTCAAACTTGTGATATCAAGCCTTAGCCTGCC 2338
DB |||||
2281 TTAAGAGCTTTTCTTAAATTTGAGAGCTCAAACTTGTGATATCAAGCCTTAGCCTGCC 2340
QY |||||
2339 TGTGAGCACTGAGCATGATCAGAGAAAGCATGGGTAAAAAGAACCAAGAGAGTCCAAAG 2398
DB |||||
2341 TGTGAGCACTGAGCATGATCAGAGAAAGCATGGGTAAAAAGAACCAAGAGAGTCCAAAG 2400
QY |||||
2399 AGCTATGCAACAAAGATCAACAGCAGTTTAAAGACCAACAGAGTAAGGACAGGCTCTGG 2458
DB |||||
2401 AGCTATGCAACAAAGATCAACAGCAGTTTAAAGACCAACAGAGTAAGGACAGGCTCTGG 2460
QY |||||
2459 AGCCCAACGAGTCCAAAGTCTTGGGTGCTGCTGAGAGTGAAGAGTCTGACACCTGG 2518
DB |||||
2461 AACCACACGAGTCCAAAGTCTTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
QY |||||
2519 GCCTCCAGCGCATCTTCTACCTGAAAGTGGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2578
DB |||||
2521 GCCTCCAGCGCATCTTCTACCTGAAAGTGGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
QY |||||
2579 ACAAGCTCTTCTGAGACACCTTACCTTCTAATCAGGAGCAGTGGAGCAGTGGAGTGCCT 2638
DB |||||
2581 ACAAGCTCTTCTGAGACACCTTACCTTCTAATCAGGAGCAGTGGAGCAGTGGAGTGCCT 2640
QY |||||
2639 CCTCTCTAGACCTGCTTGTACCGCAAGAGGATAGTTTGGAAACCTATCATTTCC 2698
DB |||||
2641 CCTCTCTAGACCTGCTTGTACCGCAAGAGGATAGTTTGGAAACCTATCATTTCC 2700
QY |||||
2699 TGTCTCTCTTAAAGAGGAAAGCAGCTCCTGTAGAAAGCAAAAGACTTTCTTTTTTCTG 2758
DB |||||
2701 TGTCTCTCTTAAAGAGGAAAGCAGCTCCTGTAGAAAGCAAAAGACTTTCTTTTTTCTG 2760
QY |||||
2759 GCTCTTTCTCTTAAACCTAAAGCAGAAACCTGACAGATTTGTTGTTGGGGTGTGTT 2818
DB |||||
2761 GCTCTTTCTCTTAAACCTAAAGCAGAAACCTGACAGATTTGTTGTTGGGGTGTGTT 2820
QY |||||
2819 TTATATTTAGGATTTGGGGATGGGGTGGGGGTTATAGTTTCATGAGGGTTTTCTAA 2878
DB |||||
2821 TTATATTTAGGATTTGGGGATGGGGTGGGGGTTATAGTTTCATGAGGGTTTTCTAA 2880
QY |||||
2879 GAAATTTGCTTAAACAAAGCAGCTTTTGGCAATGCTATCCAGCAGGAAAAAAGAGTAATA 2938
DB |||||
2881 GAAATTTGCTTAAACAAAGCAGCTTTTGGCAATGCTATCCAGCAGGAAAAAAGAGTAATA 2940
QY |||||
2939 TAACTGTTTTTAAACCTCTTTCTGGGGAAATCCAAATTAATAGTTGCTTTGTTATTTAAAAACAA 2998

CC proteins: NOV1a-1t show homology to fibroblast growth factor receptor 4
CC (FGR4); NOV2a shows homology to complement factor I precursor; NOV3a
CC shows homology to matrix metalloproteinase-15 precursor; NOV4a shows
CC homology to MDC3; NOV5a-5c show homology to T-lymphocyte surface antigen
CC Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21
CC (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like
CC polypeptide variant; NOV8a-8g show homology to antileukoprotease 1
CC precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology
CC to nuclear hormone receptor NOR-1; NOV11a-11j show homology to
CC transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin
CC -dynorphin precursor. The present sequence encodes a NOVX polypeptide of
CC the invention.
XX
SQ

Sequence 3802 BP; 919 A; 1141 C; 912 G; 830 T; 0 U; 0 Other;

Query Match 98.1%; Score 3722; DB 12; Length 3802;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY	1	ATAAATGACGTGCGGAGAGCGGAAAGCGGAGCGGAGCGGAGCGGAGTCTCTGCT	60
DB	1	ATAAATGACGTGCGGAGAGCGGAAAGCGGAGCGGAGCGGAGCGGAGTCTCTGCT	60
QY	61	CCGCGCCCGCCACCCCTCCAGCTCCTGCTCTCTCCGCTCCCATACACAGCGCTCA	120
DB	61	CCGCGCCCGCCACCCCTCCAGCTCCTGCTCTCTCCGCTCCCATACACAGCGCTCA	120
QY	121	CACCGCGCTCCCTCACTCGCACACACAGACAAAGCGCGCACAGGCTCCG--CACAC	178
DB	121	CACCGCGCTCCCTCACTCGAACACACAGACAAAGCGCGCACAGGCTCCGCGACAC	180
QY	179	ACTTCGCTCTCCGCGCGCTCACACCCCTCTGCTCTGCTGAGCGCTTCCGCGTCA	238
DB	181	ACTTCGCTCTCCGCGCGCTCACACCCCTCTGCTGAGCGCTTCCGCGTCA	240
QY	239	CGCGCGAGCTGACGCGCCCTCCGCGGCTCACTTTGCAAGCTGAGCGTCCGCGTGC	298
DB	241	CGCGCGAGCTGAGCGCCCTCCGCGGCTCACTTTGCAAGCTGAGCGTCCGCGTGC	300
QY	299	CGTGGAGTGGGAAACAGCGCGGCACTCTCCCTCTGGTACAGCCCAAGCAGGACGC	358
DB	301	CGTGGAGTGGGAAACAGCGCGGCACTCTCCCTCTGGTACAGCCCAAGCAGGACGC	360
QY	359	CGGGAACCTCTCGGCTGCTCTCCATGAGTCGGGATCGAGCATCCCGCACAGCGG	418
DB	361	CGGGAACCTCTCGGCTGCTCTCCATGAGTCGGGATCGAGCATCCCGCACAGCGG	420
QY	419	CTCACCGCTCCGCGAGCGCTGGGCTGTACACCGCAGCCCTTCGGGACAGCAGCTGT	478
DB	421	CTCACCGCTCCGCGAGCGCTGGGCTGTACACCGCAGCCCTTCGGGACAGCAGCTGT	480
QY	479	GACTCCCGCCAGTGAGATTTGGGACAGCTCTCTAGAAACTCTGCTCTAAAGACGGAAC	538
DB	481	GACTCCCGCCAGTGAGATTTGGGACAGCTCTCTAGAAACTCTGCTCTAAAGACGGAAC	540
QY	539	CGGCAAGCACTCAAAAGCCACTGCGGAAAGAGGCGAGCGCGGAGCGGCGCTGAGC	598
DB	541	CGGCAAGCACTCAAAAGCCACTGCGGAAAGAGGCGAGCGCGGAGCGGCGCTGAGC	600
QY	599	CTGGACCCCTAGCGGTGCGGCGAGCACTGCGGCGCTTTCGCTCGCGGAGCTCGGCTC	658
DB	601	CTGGACCCCTAGCGGTGCGGCGAGCACTGCGGCGCTTTCGCTCGCGGAGCTCGGCTC	660
QY	659	CTCCTACACTCTCAGCGCTCGGTGAGAGACCCCGAGCCCGCCACTTCAGCGCGAAGAT	718
DB	661	CTCCTACACTCTCAGCGCTCGGTGAGAGACCCCGAGCCCGCCACTTCAGCGCGAAGAT	720
QY	719	ACCCTCCAGATAGCCCTCGCTCAAGCCCAATATAGCCCTTCCCTCCAGTTCCAGTT	778
DB	721	ACCCTCCAGATAGCCCTCGCTCAAGCCCAATATAGCCCTTCCCTCCAGTTCCAGTT	780
QY	779	ATGCGGCGGACATACAGCTGGGAATACACAGCGAGATCATGAACCCCGCATACCA	838

DB	781	ATGCGGCGGACATACAGCTCGGAATACACACGAGAGATCATGAACCCCGACTACACCA	840
QY	839	AGCTGACCAATGAGACCTTGGCAGCACTGAGATCAGGCTACAGCCACACGTCCTGCCCA	898
DB	841	AGCTGACCAATGAGACCTTGGCAGCACTGAGATCAGGCTACAGCCACACGTCCTGCCCA	900
QY	899	GCATCAGTACCTTCGTGGAGGCTACTCGAGCAACTACCAACTCAAGCCCTTCGCGTGT	958
DB	901	GCATCAGTACCTTCGTGGAGGCTACTCGAGCAACTACCAACTCAAGCCCTTCGCGTGT	960
QY	959	ACCAATGACGCGGCGCTTGGATCAAGTGGAGGAGGCGGCGCGCCAGCTACCATCACC	1018
DB	961	ACCAATGACGCGGCGCTTGGATCAAGTGGAGGAGGCGGCGCGCCAGCTACCATCACC	1020
QY	1019	ATCAC	1078
DB	1021	ATCAC	1080
QY	1079	TTCTCTCCAGCTCCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1138
DB	1081	TTCTCTCCAGCTCCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1140
QY	1139	AGTCCCGCCACCGTCCACCGCCACACACCGCGCTTCCCGCGGAGGCGGCGGCTTATGG	1198
DB	1141	AGTCCCGCCACCGTCCACCGCCACACACCGCGCTTCCCGCGGAGGCGGCGGCTTATGG	1200
QY	1199	ACGAGGACTGCGCTCCGCGCGCGCTGCATCGCACCGCGCGCTGCTGGACCGCGCGCA	1258
DB	1201	ACGAGGACTGCGCTCCGCGCGCGCTGCATCGCACCGCGCGCTGCTGGACCGCGCGCA	1260
QY	1259	TGAAGCGGCTCCGACGCTGGCGGCGCGCTTCCCGCTCTTCCACTTCAAGCCCTCCG	1318
DB	1261	TGAAGCGGCTCCGACGCTGGCGGCGCGCTTCCCGCTCTTCCACTTCAAGCCCTCCG	1320
QY	1319	CGCGCATCCCCCG	1378
DB	1321	CGCGCATCCCCCG	1380
QY	1379	CGCTGCGCGCTCAGCTGCGCTGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1438
DB	1381	CGCTGCGCGCTCAGCTGCGCTGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1440
QY	1439	CGCTTGAGACCAACCGGTACGCGCTGGCGTGGCGCAAGAGGCGGCGCGCGCTTCC	1498
DB	1441	CGCTTGAGGCGCAACCGGTACGCGCTGGCGTGGCGCAAGAGGCGGCGCGCGCTTCC	1500
QY	1499	CGCGCTCTCGGCTCAGCGCTTCCCTTACCGCGTCCAGCTGCTGGCGAGAGTCCAGCG	1558
DB	1501	CGCGCTCTCGGCTCAGCGCTTCCCTTACCGCGTCCAGCTGCTGGCGAGAGTCCAGCG	1560
QY	1559	TGCGCTCG	1618
DB	1561	TGCGCTCG	1620
QY	1619	ACAAACCGCGCTCG	1678
DB	1621	ACAAACCGCGCTCG	1680
QY	1679	AGAGAACAGTGCAGAAATATGTTGCTGCGCAATATAAACTCCCGAGTAG	1738
DB	1681	AGAGAACAGTGCAGAAATATGTTGCTGCGCAATATAAACTCCCGAGTAG	1740
QY	1739	ACAAGAGAGTGCAGAAATATGTTGCTGCGCAATATAAACTCCCGAGTAG	1798
DB	1741	ACAAGAGAGTGCAGAAATATGTTGCTGCGCAATATAAACTCCCGAGTAG	1800
QY	1799	TGTTAAAGAGTTCGCGTACAGATAGTCTGAAAGGAGGAGAGGTCGTCTGCGCTTCA	1858
DB	1801	TGTTAAAGAGTTCGCGTACAGATAGTCTGAAAGGAGGAGAGGTCGTCTGCGCTTCA	1860
QY	1859	RACCAAGAGGCGCATTAACAGGAACTTCTCAGCGCTCTCCACTTCTCTCCATCT	1918
DB	1861	AACCAAGAGGCGCATTAACAGGAACTTCTCAGCGCTCTCCACTTCTCTCCATCT	1920

QY	1919	GCATGATGAATGCCCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTATT	1978
Db	1921	GCATGATGAATGCTCTTGTGCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTATT	1980
QY	1979	CCAGATCTGTCCCACTGACAGGCTGTGAGGACACAGATGCTGAGCATGTGCAACAAT	2038
Db	1981	CCAGATCTGTCCCACTGACAGGCTGTGAGGACACAGATGCTGAGCATGTGCAACAAT	2040
QY	2039	TCATAAAGCTCTGACAGCTCCATTGATGATATCCAGAGCTGCGCAGAAAGATTCCGG	2098
Db	2041	TCATAAAGCTCTGACAGCTCCATTGATGATATCCAGAGCTGCGCAGAAAGATTCCGG	2100
QY	2099	GATTTACTGATCTCCCAAGAAAGATCAGACATTAATTATGAATCAGCTTTTGGAGC	2158
Db	2101	GATTTACTGATCTCCCAAGAAAGATCAGACATTAATTATGAATCAGCTTTTGGAGC	2160
QY	2159	TGTTGTCTCTGATCTTTCATCATGAGTCAAAACACTGCTGAAGATTAAGTTGTCTGCA	2218
Db	2161	TGTTGTCTCTGATCTTTCATCATGAGTCAAAACACTGCTGAAGATTAAGTTGTCTGCA	2220
QY	2219	ATGGACTTGTCTGATCGACTTCAGTCTGCTTGTGGATTTGGGAGTGGCTCGACTCTA	2278
Db	2221	ATGGACTTGTCTGATCGACTTCAGTCTGCTTGTGGATTTGGGAGTGGCTCGACTCTA	2280
QY	2279	TTAAAGACTTTTCTTAAATTTGCAGAGCTGAACTTTGATATCAAGCTTTAGCTGCC	2338
Db	2281	TTAAAGACTTTTCTTAAATTTGCAGAGCTGAACTTTGATATCAAGCTTTAGCTGCC	2340
QY	2339	TGTCAGACTGAGCATGATCAGAGAAAGACATGGGTTAAAGAACCAAGAGAGTCGAAG	2398
Db	2341	TGTCAGACTGAGCATGATCAGAGAAAGACATGGGTTAAAGAACCAAGAGAGTCGAAG	2400
QY	2399	AGCTATGCAACAAAGATCAACAGCAGTTTAAAGACACACAGAGTAAGGACAGGCTCTGG	2458
Db	2401	AGCTATGCAACAAAGATCAACAGCAGTTTAAAGACACACAGAGTAAGGACAGGCTCTGG	2460
QY	2459	AGCCCAACGAGTCCAAAGTCTCTGGTGCCTTGTAGAACTGAGGAAGATCTGACCCCTGG	2518
Db	2461	AGCCCAACGAGTCCAAAGTCTCTGGTGCCTTGTAGAACTGAGGAAGATCTGACCCCTGG	2520
QY	2519	GGCTCAGGCGATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATG	2578
Db	2521	GGCTCAGGCGATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATG	2580
QY	2579	ACAAAGCTCTTCTGACACCTTACCTTCTAATCAGGAGCAGTGGAGCAGTGGCTGCT	2638
Db	2581	ACAAAGCTCTTCTGACACCTTACCTTCTAATCAGGAGCAGTGGAGCAGTGGCTGCT	2640
QY	2639	CCTCTCCTAGCACCTTGTCTACGCAAGAGGATAGGTTTGGAAACCTTATCATTTCC	2698
Db	2641	CCTCTCCTAGCACCTTGTCTACGCAAGAGGATAGGTTTGGAAACCTTATCATTTCC	2700
QY	2699	TGTCCTTCTTAAAGAGAAACAGCTCTCTGTAAGCAAGACATTTCTTTTTTCTG	2758
Db	2701	TGTCCTTCTTAAAGAGAAACAGCTCTCTGTAAGCAAGACATTTCTTTTTTCTG	2760
QY	2759	GGCTCTTTCTTACAACTTAAAGCAGAAAACTGCGAGAGTATGTGTGGGTTGTGT	2818
Db	2761	GGCTCTTTCTTACAACTTAAAGCAGAAAACTTGCAGAGTATGTGTGGGTTGTGT	2820
QY	2819	TTATATTTTAGGCATTTGGGGATGGGGTGGAGGGGTTATAGTTTATGAGGGTTTCTAA	2878
Db	2821	TTATATTTTAGGCATTTGGGGATGGGGTGGAGGGGTTATAGTTTATGAGGGTTTCTAA	2880
QY	2879	GAATTTGCTAAACAAAGCATTTTGGACATGCTATCCAGCAGGAAAGAAAGGATATA	2938
Db	2881	GAATTTGCTAAACAAAGCATTTTGGACATGCTATCCAGCAGGAAAGAAAGGATATA	2940
QY	2939	TAACTGTTTTTAAACCTTTCTTCTGGGAAATCCAAATTATAGTTTGTATTTAAAAACAA	2998
Db	2941	TAACTGTTTTTAAACCTTTCTTCTGGGAAATCCAAATTATAGTTTGTATTTAAAAACAA	3000

QY	2999	GAAACAGCAAGGTTGTTCCGACAGGTTAGATGTCTTAAAGATTGTCCTCTGAAAAAT	3058
Db	3001	GAAACAGCAAGGTTGTTCCGACAGGTTAGATGTCTTAAAGATTGTCCTCTGAAAAAT	3060
QY	3059	ATGCTTCTCTGATCAAAAGGTACGTTATGTTGGTGCAAAACAGGCAGAACTTCTTTTAAT	3118
Db	3061	ATGCTTCTCTGATCAAAAGGTACGTTATGTTGGTGCAAAACAGGCAGAACTTCTTTTAAT	3120
QY	3119	TCCTTCTTCTCTTATTTTAAACAAATGGTGAAGATTGAGGATTACCTACAAATCAGACAT	3178
Db	3121	TCCTTCTTCTCTTATTTTAAACAAATGGTGAAGATTGAGGATTACCTACAAATCAGACAT	3180
QY	3179	GGCAAAAACAAATATATGGCTGTTTGTTCATATAACAAAGTGCATTTTAAAGTGTCT	3238
Db	3181	GGCAAAAACAAATATATGGCTGTTTGTTCATATAACAAAGTGCATTTTAAAGTGTCT	3240
QY	3239	TACTAAGTCTTGTATTATTAACCTCTCTTATTTCTATATGGAATAAAAAGGAGGAGTCA	3298
Db	3241	TACTAAGTCTTGTATTATTAACCTCTCTTATTTCTATATGGAATAAAAAGGAGGAGTCA	3300
QY	3299	TGTTAGCAAAATGACACAGTTAATATCCCTAGCAGAGGCTGTGTTCACCTTCCCTGTGATC	3358
Db	3301	TGTTAGCAAAATGACACAGTTAATATCCCTAGCAGAGGCTGTGTTCACCTTCCCTGTGATC	3360
QY	3359	CCTTCTGAGGTATGGCCCATCCAAAGCTTTTATAGGCCATTTCTGTATGGAACAGATCCCTG	3418
Db	3361	CCTTCTGAGGTATGGCCCATCCAAAGCTTTTATAGGCCATTTCTGTATGGAACAGATCCCTG	3420
QY	3419	CCCTGACGTCCAGCTATCCCTGAAAGTGGATCAGATTAATACTGGATTACATGTAACCTG	3478
Db	3421	CCCTGACGTCCAGCTATCCCTGAAAGTGGATCAGATTAATACTGGATTACATGTAACCTG	3480
QY	3479	TTTTGGTTGTCTTCTATCAACCCCAACAGAGTTCCTTAAACTTGTCTCAGTTATAGTAAC	3538
Db	3481	TTTTGGTTGTCTTCTATCAACCCCAACAGAGTTCCTTAAACTTGTCTCAGTTATAGTAAC	3540
QY	3539	TGACTGTGATATTCATTCAGAGGCCCATAGTCTGATGATTTGATCCCTAGATAAG	3598
Db	3541	TGACTGTGATATTCATTCAGAGGCCCATAGTCTGATGATTTGATCCCTAGATAAG	3600
QY	3599	AACATGCAAAATCAGCAGGAACTGTGTATACAGGGTAAAGCAGGAGCAATAAGGATTTT	3658
Db	3601	AACATGCAAAATCAGCAGGAACTGTGTATACAGGGTAAAGCAGGAGCAATAAGGATTTT	3660
QY	3659	TATAGATATAATTTAAATTTTGTTA--TTGGTTAAGGAGA--CAATTTTGGAGAGCAAGCA	3715
Db	3661	TATAGATATAATTTAAATTTTGTTA--TTGGTTAAGGAGA--CAATTTTGGAGAGCAAGCA	3720
QY	3716	AA---TCCTTTTAAAAATAGTATGAATGTGAATAGTAAAGATTTTAAAAAATAGTAT	3772
Db	3721	AATCTTCTTTTAAAAAATAGTATGAATGTGAATAGTAAAGATTTTAAAAAATAGTAT	3774
QY	3773	GAGTGTGAGTACTAGGAAGGAT 3794	
Db	3781	GAGTGTGAGTACTAGGAAGGAT 3802	

RESULT 5

ADE25701

ID ADE25701 standard; cDNA; 5828 BP.

3541	ACTGACTGGTATATCAATTCAGAAAGCCCAATAGTCACTAGTTAGTATTTGTATTCCTTAGATA	3600
Db		
3597	AGAACTGCAAAATCAGCAGGAACTGGTTCATACAGGGTAAAGCACCAGGAGCAATAAAGGATT	3656
Qy		
3601	AGAACTGCAAAATCAGCAGGAACTGGTTCATACAGGGTAAAGCACCAGGAGCAATAAAGGATT	3660
Db		
3657	TTTATAGATATAATTTTAAATTTTGGTTATTCGGTTAAAGGAGACAAATTTTGGAGAGCAAGCAA	3716
Qy		
3661	TTTATAGATATAATTTTAAATTTTGGTTATTCGGTTAAAGGAGACAAATTTTGGAGAGCAAGCAA	3720
Db		
3717	A---TCCTTTTAAAAAATAGTATGCAATGTCGAATACTAGAAAGAAATTTAAAAAATATGATATG	3773
Qy		
3721	ATCTTCTTTTTAAAAAATAGTATGCAATGTCGAATACTAGAAAGAAATTTAAAGAAATAGTATG	3780
Db		
3774	AGTGTGAGTACTAGGAAGGAT	3794
Qy		
3781	AGTGTGAGTACTAGGAAGGAT	3801
Db		

RESULT 6

RESULTS 6
ABX62937
ID ABX62937 standard; cDNA; 5642 BP.

AC ABX62937;

DT 25-FEB-2003 (first entry)

Human activated T cell cDNA #53.

T cell; gene; ss; differential expression; T cell activation;
 KW: anti-allergic; cytostatic; immunosuppressive; antimicrobial; gene therapy;
 KW: allergy; cancer; graft versus host disease; infection;
 KW: autoimmune disorder.

OS Homo sapiens.

AA
PN
US2002137077-A1.

PD 26-SEP-2002.

PF 25-OCT-2001; 2001US-00002600.

PR 25-OCT-2000; 2000US-0243521P.

РА (НОРК/) НОРКИНС С М.

PA (PETE/) PETERSON D P.

PA (COCK/) COCKS B G.

PA (HAWK/) HAWKINS P R.

PI Hopkins CM, Peterson DP, Cocks BG, Hawkins PR;

DR WPI; 2003-102381/09.

PT New combination comprising several cDNAs that are differentially
PT expressed in activated T cells, useful for diagnosing, treating, staging
PT or monitoring treatment for allergy, cancer, infectious and/or autoimmune
PT disorders.

PS Claim 1: Page: 180pp: English.

This invention relates to the sequences of several cDNAs that are differentially expressed in activated T cells. The sequences of the invention may have antiallergic, cytostatic, immunosuppressive and antimicrobial activity and may be used in gene therapy. The invention also comprises a method for screening samples for differentially expressed genes and a method for detecting these cDNAs by hybridisation. The methods and compositions of the present invention are useful for: diagnosing, treating, staging or monitoring treatment for allergy; cancer, chronic graft versus host disease, infectious and/or autoimmune disorders. The present sequence represents a cDNA of the invention that is differentially expressed in activated T cells

SQ Sequence 5642 BP; 1472 A; 1399 C; 1212 G; 1465 T; 0 U; 94 Other;

Query Match	96.7%	Score 3668.2	DB 8	Length 5642
Best Local Similarity	98.1%	Pred. No. 0		
Matches 3728	Conservative 0	Mismatches 66	Indels 8	Gaps 5
Qy	1	ATAAATGAGTGC	CGGAGAGCGAGCGGAAAGCGCGACGCGGAGAGCGGAGTCTCCTGCCT	60
Db	1	ATAAATGAGTGC	CGGAGAGCGAGCGGAAAGCGCGACGCGGAGAGCGGAGTCTCCTGCCT	60
Qy	61	CCGCGCCCCAC	CCCTCCAGTCTCTCTCTCCGCTCCCATACACAGACGCGCTCA	120
Db	61	CCGCGCCCCAC	CCCTCCAGTCTCTCTCTCCGCTCCCATACACAGACGCGCTCA	120
Qy	121	CACCGCTCTCT	CACTCGCACACACAGAACAGACGAGCTCG--CACACAC	178
Db	121	CACCGCTCTCT	CACTCGAACACACAGAACAGACGAGCTCCGACACACAC	180
Qy	179	ACTTCGCTCT	CCCGCGGCTCACACCCCTTTGCTGAGCCCTTCCCGGTGACGCGG	238
Db	181	ACTTCGCTCT	CCCGCGGCTCACACCCCTTTGCTGAGCCCTTCCCGGTGACGCGCG	240
Qy	239	CGCGCAGCTG	GACGCGCTCCGCGGCTCACTTTGCAACGCTGACGGTCCGCGACGTGGC	298
Db	241	CGCGCAGCTG	GACGCGCTCCGCGGCTCACTTTGCAACGCTGACGGTCCGCGACGTGGC	300
Qy	299	CGTGGAGTGG	AAACAGCGCGCGCATCTCCCCCTTGGTCACAGCCCAAGCCAGGACGCC	358
Db	301	CGTGGAGTGG	AAACAGCGCGCGCATCTCCCCCTTGGTCACAGCCCAAGCCAGGACGCC	360
Qy	359	CGCGGAACCT	CTCTCGGCTGTGCTCTCCCATGAGTGGGATCGAGATCCCCACACGCGG	418
Db	361	CGCGGAACCT	CTCTCGGCTGTGCTCTCCCATGAGTGGGATCGAGATCCCCACACGCGG	420
Qy	419	CTCACCGCT	CCGGGAGCGCTGGGCTGTACACCGCAGCCTTCCGGGACAGCAGCTGT	478
Db	421	CTCACCGCT	CCGGGAGCGCTGGGCTGTGTACACCGCAGCCTTCCGGGACAGCAGCTGT	480
Qy	479	GACTCCCCC	CCAGTGAGATTTCCGGACAGCTCTCTAGAAACTCGCTCTAAAGACGGAAC	538
Db	481	GACTCCCCC	CCAGTGAGATTTCCGGGACAGCTCTCTAGAAACTCGCTCTAAAGACGGAAC	540
Qy	539	CGCCACAGCA	CTCAAGCCCACTGCGGGAAGAGGGCAGCCCGGGAAGCCCGGCGCTGAGC	598
Db	541	CGCCACAGCA	CTCAAGCCCACTGCGGGAAGAGGGCAGCCCGGGAAGCCCGGCGCTGAGC	600
Qy	599	CTGGACCTCT	TAGCGGTGCGGGCAGACTGCGGGCGCTTCGCTCGCGGAGCTCCGCTC	658
Db	601	CTGGACCTCT	TAGCGGTGCGGGCAGACTGCGGGCGCTTCGCTCGCGGAGCTCCGCTC	660
Qy	659	CTCCTACACT	CTCAGGCTCCGCTGGAGAGACCCCGAGCCCCACATTCAGCGCGCAAGAT	718
Db	661	CTCCTACACT	CTCAGGCTCCGCTGGAGAGACCCCGAGCCCCACATTCAGCGCGCAAGAT	720
Qy	719	ACCCCTCCAG	ATATGCCCTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT	778
Db	721	ACCCCTCCAG	ATATGCCCTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT	780
Qy	779	ATCGCGGGCA	GACATACAGCTCGGAATACACACGAGATCATGAACCCCGACTACACCA	838
Db	781	ATCGCGGGCA	GACATACAGCTCGGAATACACACGAGATCATGAACCCCGACTACACCA	840
Qy	839	AGCTGACCAT	TGGACCTTGGCAGCCTAGATACGGCTACAGCCACACGCTCCCTGCCCA	898
Db	841	AGCTGACCAT	TGGACCTTGGCAGCCTAGATACGGCTACAGCCACACGCTCCCTGCCCA	900
Qy	899	GCATCAGTAC	CTTCTGTGGAGGGCTACTCGAGCAACTACGAACTCAAGCCCTTCTCGGTGT	958
Db	901	GCATCAGTAC	CTTCTGTGGAGGGCTACTCGAGCAACTACGAACTCAAGCCCTTCTCGGTGT	960
Qy	959	ACCAATGCA	GCGGCTTGATCAAGTGGAGAGGGCGGGCGCCAGCTACCATCACC	1018
Db	961	ACCAATGCA	GCGGCTTGATCAAGTGGAGAGGGCGGGCGCCAGCTNNNNNNNN	1020

Db 3181 CATGCAAAACAATAATGCTGTTGCTTCCATAAACAAGTGCAATTTTTTAAAGTGCTG 3240
 Qy 3236 TCTTACTAAGTCTTGTTTATTAATCTCTCTTTATCTATATGGAATAAAGAGGCGAG 3295
 Db 3241 TCTTACTAAGTCTTGTTTATTAATCTCTCTTTATCTATATGGAATAAAGAGGCGAG 3300
 Qy 3296 TCATGTTAGCAATGACAGTGAATTAATCTCTAGCAGAGCTGTGTTTACCTTCCCTGTGCG 3355
 Db 3301 TCATGTTAGCAATGACAGTGAATTAATCTCTAGCAGAGCTGTGTTTACCTTCCCTGTGCG 3360
 Qy 3356 ATCCCTCTTGAGGTATGGCCCATCCAAAGACTTTTAGGCCAATCTTGTATGGAACCCAGATCC 3415
 Db 3361 ATCCCTCTTGAGGTATGGCCCATCCAAAGACTTTTAGGCCAATCTTGTATGGAACCCAGATCC 3420
 Qy 3416 CTGCCCTGACTGCTCAGCTATCTCTGAAAGTGATCAGATTATAAAGTGAATACATGTAA 3475
 Db 3421 CTGCCCTGACTGCTCAGCTATCTCTGAAAGTGATCAGATTATAAAGTGAATACATGTAA 3480
 Qy 3476 CTGTTTGGTGTGTTTCTATCAACCCACAGAGTTCCCTAAACTTGTCTTCAAGTTATAGT 3535
 Db 3481 CTGTTTGGTGTGTTTCTATCAACCCACAGAGTTCCCTAAACTTGTCTTCAAGTTATAGT 3540
 Qy 3536 AACTGACTGCTATATTCATTTCAGAAGCGCCATAAGTTCAGTTGATTTTATCCCTAGAT 3595
 Db 3541 AACTGACTGCTATATTCATTTCAGAAGCGCCATAAGTTCAGTTGATTTTATCCCTAGAT 3600
 Qy 3596 AAGAACATCAATCAGCAGGAAGCTGTCATCAGGGTAAGCACAAGGAGCAATTAAGGAT 3655
 Db 3601 AAGAACATCAATCAGCAGGAAGCTGTCATCAGGGTAAGCACAAGGAGCAATTAAGGAT 3660
 Qy 3656 TTTTATAGATATAATTTAATTTTGTATTTGTTTATGTTAAGGAGACAATTTTGGAGAGAAGCA 3715
 Db 3661 TTTTATAGATATAATTTAATTTTGTATTTGTTTATGTTTAAAGGAGACAATTTTGGAGAGAAGCA 3720
 Qy 3716 AA---TCTTTTAAATAATAGTATGAATGTGAATCTAGAAAGATTTTAAATAATAGTAT 3772
 Db 3721 AATCTCTTTTAAATAATAGTATGAATGTGAATCTAGAAAGATTTTAAATAATAGTAT 3780
 Qy 3773 GAGTGTGAGTACTAGGAAGGAT 3794
 Db 3781 GAGTGTGAGTACTAGGAAGGAT 3802

RESULT 7

AAT73334
 ID AAT73334 standard; cDNA to mRNA; 3762 BP.

AC AAT73334;

XX 08-SEP-1997 (first entry)

DE Human neuron-derived orphan receptor NOR-1 gene.

DE Human; neuron-derived orphan receptor; open reading frame; homology; rat;
 KW NOR-1; DNA binding domain; brain; neuronal; Alzheimer's disease;
 KW Parkinson's disease; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 PH 692..2572
 FT /*tag= a
 FT /product= "NOR-1 orphan receptor"

FT JP09084585-A.

PN 31-MAR-1997.

PD 21-SEP-1995; 95JP-00242909.

PF 21-SEP-1995; 95JP-00242909.

PR 21-SEP-1995; 95JP-00242909.

XX (TERU) TERUMO CORP.

XX WPI; 1997-253003/23.
 DR P-PSDB; AAW16398.
 XX Human neuron-derived orphan nuclear receptor gene - useful for diagnosing
 PT brain diseases such as Alzheimer's or Parkinson's disease or
 PT neuroblastoma.
 XX Claim 1; Page 5-8; 8pp; Japanese.
 XX This is the nucleotide sequence of a novel human neuron-derived orphan
 CC nuclear receptor gene. The open reading frame encodes a protein of 626
 CC amino acids, which has high homology to the rat NOR-1 receptor. The
 CC protein contains several DNA binding domains including P, D and A boxes
 CC which also have high homology to the rat NOR-1 DNA binding domains. The
 CC nucleic acid and protein can be used to diagnose brain neuronal diseases,
 CC such as Alzheimer's disease, Parkinson's disease. Further, the genetic
 CC diseases can be treated using the receptor as it inhibits expression of a
 CC gene by binding to it
 XX Sequence 3762 BP; 907 A; 1128 C; 898 G; 829 T; 0 U; 0 Other;
 SQ Query Match 96.3%; Score 3653.6; DB 2; Length 3762;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 3728; Conservative 0; Mismatches 24; Indels 10; Gaps 5;

Qy 43 GAGCGGAGTCTCTGCTCCGCGCCGCCACCCCTCCAGCTCTGCTCTCTCCGCTCCC 102
 Db 1 GAGCGGAGTCTCTGCTCCGCGCCGCCACCCCTCCAGCTCTGCTCTCTCCGCTCCC 60
 Qy 103 CATACACAGACGCGCTCACAACCGCTCCCTCACTCGCACACACAGACACAAGGCGCAC 162
 Db 61 CATACACAGACGCGCTCACAACCGCTCCCTCACTCGCACACACAGACACAAGGCGCAC 120
 Qy 163 CAGGCTCCG--CACACACACTTTCGCTCTCCCGCGGCTCACACCCCTCTTGGCCTGAGCC 220
 Db 121 CAGGCTCCGCGCACACACACTTTCGCTCTCCCGCGGCTCACACCCCTCTTGGCCTGAGCC 180
 Qy 221 CTTGCGGCTGACGCGCGCGCGAGCTGGAGCGCCCTCCCGGGTCACTTTTGCAACGCT 280
 Db 181 CTTGCGGCTGACGCGCGCGCGAGCTGGAGCGCCCTCCCGGGTCACTTTTGCAACGCT 240
 Qy 281 GACGCTGCGGAGTGGCGGTGGAGTGGGAACAGCGGGGCGATCTCCCGCTTGGTCTAC 340
 Db 241 GACGCTGCGGAGTGGCGGTGGAGTGGGAACAGCGGGGCGATCTCTCCCGCTTGGTCTAC 300
 Qy 341 AGCCCAA--GCCAGGACGCCCGCGAACTCTCGGCTGTGCTCTCCCATGAGTCCGGATC 398
 Db 301 AGCCCAAAGCCAGGAGCGCCCGGAACTCTCGGCTGTGCTCTCCCATGAGTCCGGATC 360
 Qy 399 GCAGCATCCCGACACGCGGCTCACCGGCTCCGGAGCGGCTGGGCTTGTACACGCGAGC 458
 Db 361 GCAGCATCCCGACACGCGGCTCACCGGCTCCGGAGCGGCTGGGCTTGTACACGCGAGC 420
 Qy 459 CTTCCGGGACAGAGCTGTGACTCCCGCCGAGTGCAGATTTCGGGACAGCTCTCTAGAA 518
 Db 421 CTTCCGGGACAGAGCTGTGACTCCCGCCGAGTGCAGATTTCGGGACAGCTCTCTAGAA 480
 Qy 519 ACTGCTCTTAAAGACGGAAACGCCACAGCACTCAAAGCCCACTGCGGAAGAGGCGAGCCC 578
 Db 481 ACTGCTCTTAAAGACGGAAACGCCACAGCACTCAAAGCCCACTGCGGAAGAGGCGAGCCC 540
 Qy 579 GCAAGCCCGGGCCCTGAGCTTGGACCTTAGCGGTGCGGGGAGCAGCTGCGCGGCGCTTC 638
 Db 541 GCAAGCCCGGGCCCTGAGCTTGGACCTTAGCGGTGCGGGGAGCAGCTGCGCGGCGCTTC 600
 Qy 639 GCCTCGCGGAGCTCCGCTCTCTTACACTCTACGCTTCGCTGAGAGAGACCCCGAGCCC 698
 Db 601 GCCTCGCGGAGCTCCGCTCTCTTACACTCTACGCTTCGCTGAGAGAGACCCCGAGCCC 660
 Qy 699 CACCATTCCAGCGCGCAAGATACCTCCAGATATGCCCTCCAGATATGCCCTCCAGCCCAATATAGCCC 758
 Db 661 CACCATTCCAGCGCGCAAGATACCTCCAGATATGCCCTCCAGATATGCCCTCCAGCCCAATATAGCCC 720

QY 759 TTCCCTCCAGTTCAGTTATGCGCGCAGACATACAGCTCGGAATACACACGAGAT 818
DB 721 TTCCCTCCAGTTCAGTTATGCGCGCAGACATACAGCTCGGAATACACACGAGAT 780
QY 819 CATGAACCCCGACTACACCAAGCTGACCATGGACCTTGGCAGCACTGAGATCAAGGCTAC 878
DB 781 CATGAACCCCGACTACACCAAGCTGACCATGGACCTTGGCAGCACTGAGATCAAGGCTAC 840
QY 879 AGCCACGAGTCCCTGCCAGCATCAGTACCTTGGTGGAGGGCTACTCGAGCACTACGA 938
DB 841 AGCCACGAGTCCCTGCCAGCATCAGTACCTTGGTGGAGGGCTACTCGAGCACTACGA 900
QY 939 ACTCAAGCTTCTCCGCTGTACCAATGACGGGCTTGTATCAAGTGGAGGAGGGCG 998
DB 901 ACTCAAGCTTCTCCGCTGTACCAATGACGGGCTTGTATCAAGTGGAGGAGGGCG 960
QY 999 GCGGCCAGCTACCATCAACATCACCAACCAACCAACCAACCAACCAACCAACCAACCA 1058
DB 961 GCGGCCAGCTACCATCAACATCACCAACCAACCAACCAACCAACCAACCAACCAACCA 1020
QY 1059 GCAGCATCAGAGCCATCCATTCCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAG 1118
DB 1021 GCAGCATCAGAGCCATCCATTCCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAG 1080
QY 1119 CACCTCCATGTACTTCAAGCAGTCCCAACCGTCCACCCGCCACCAACCGCGGCTTCCGCC 1178
DB 1081 CACCTCCATGTACTTCAAGCAGTCCCAACCGTCCACCCGCCACCAACCGCGGCTTCCGCC 1140
QY 1179 GCAGCGGGGGGTTATGGGACGAGGCACTGCTCCGCGGCGCGCTGCATCGCAACCGG 1238
DB 1141 GCAGCGGGGGGTTATGGGACGAGGCACTGCTCCGCGGCGCGCTGCATCGCAACCGG 1200
QY 1239 CCGCTGCTGACCGCGCGATGAAGCGGTCCCAACGCTGCGCGCGCGCGCTTCCCGCT 1298
DB 1201 CCGCTGCTGACCGCGCGATGAAGCGGTCCCAACGCTGCGCGCGCGCGCTTCCCGCT 1260
QY 1299 CTTCCACTTCAAGCCCTCGCGCGCATCCCGCGCGCATCCCGCGCGCGCGCGCGCA 1358
DB 1261 CTTCCACTTCAAGCCCTCGCGCGCATCCCGCGCGCATCCCGCGCGCGCGCGCGCA 1320
QY 1359 CTTCCACTTCAAGCCCTCGCGCGCATCCCGCGCGCATCCCGCGCGCGCGCGCGCA 1418
DB 1321 CTTCCACTTCAAGCCCTCGCGCGCATCCCGCGCGCATCCCGCGCGCGCGCGCGCA 1380
QY 1419 CCGCGCGCGCGCGCGCGCTTGAAGCCACCGCTGAGGCGCGCGCTGCGCGCGCGCGCA 1478
DB 1381 CCGCGCGCGCGCGCGCGCTTGAAGCCACCGCTGAGGCGCGCGCTGCGCGCGCGCGCA 1440
QY 1479 GCGCGCGCGCGCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCC 1538
DB 1441 GCGCGCGCGCGCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCC 1500
QY 1539 GCGCGCGCGCGCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCC 1598
DB 1501 GCGCGCGCGCGCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCC 1560
QY 1599 CAGCTGCGCGGTGCGGGGCAACGCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCC 1658
DB 1561 CAGCTGCGCGGTGCGGGGCAACGCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCC 1620
QY 1659 GCGCTGCAAGGGCTTTTCAAGAGAACAGTGCAGAGAAATGCAAAATATGTTTGCCTGCG 1718
DB 1621 GCGCTGCAAGGGCTTTTCAAGAGAACAGTGCAGAGAAATGCAAAATATGTTTGCCTGCG 1680
QY 1719 AAATAAAATCTGCCAGTAGACAAGAGCTGCAAAACCGATGTCTAGTCTGCAATTTCA 1778
DB 1681 AAATAAAATCTGCCAGTAGACAAGAGCTGCAAAACCGATGTCTAGTCTGCAATTTCA 1740
QY 1779 GAAGTGTCTCAGTGTGGAAATGCTAAAGAGTGTCCGTACAGATAGTCTGAAAGGGAG 1838
DB 1741 GAAGTGTCTCAGTGTGGAAATGCTAAAGAGTGTCCGTACAGATAGTCTGAAAGGGAG 1800

QY 1839 GAGAGGTGCTGCTGCTTCCAAACCAAGAGCCCAATTACAAAGAGAACTTCTCAGCCCTC 1898
DB 1801 GAGAGGTGCTGCTGCTTCCAAACCAAGAGCCCAATTACAAAGAGAACTTCTCAGCCCTC 1860
QY 1899 TCCACCTTCTCCTCCAAATCTGCATGATGAATGCCCTTGTCCGAGCTTTAAAGAGCTCAAC 1958
DB 1861 TCCACCTTCTCCTCCAAATCTGCATGATGAATGCCCTTGTCCGAGCTTTAAAGAGCTCAAC 1920
QY 1959 ACCCAGAGATCTTCATTTATCCAGATAGTGTCCCACTGACAGGCTGCTGAGGACACAGA 2018
DB 1921 ACCCAGAGATCTTCATTTATCCAGATAGTGTCCCACTGACAGGCTGCTGAGGACACAGA 1980
QY 2019 TGCTGAGCATGTGCAACAATTTCTAAACCTCTGACAGCTTCCATTTGATGATATCCAGAAG 2078
DB 1981 TGCTGAGCATGTGCAACAATTTCTAAACCTCTGACAGCTTCCATTTGATGATATCCAGAAG 2040
QY 2079 CTGGGCAAGAAAGATTTGGGGATTTACTGATCTCCCAAGAAAGATCAGACATTTACTTTAT 2138
DB 2041 CTGGGCAAGAAAGATTTGGGGATTTACTGATCTCCCAAGAAAGATCAGACATTTACTTTAT 2100
QY 2139 TGAATCAGCCCTTTTGGAGCTGTTTGTCTCAGACTTTTCCATCAGGTCAAAACACTGCTGA 2198
DB 2101 TGAATCAGCCCTTTTGGAGCTGTTTGTCTCAGACTTTTCCATCAGGTCAAAACACTGCTGA 2160
QY 2199 AGATAAGTTTGTGTTCTGCAATGGACTTGTCTCAGATCGACTTCACTGCTTCTGCGGAT 2258
DB 2161 AGATAAGTTTGTGTTCTGCAATGGACTTGTCTCAGATCGACTTCACTGCTTCTGCGGAT 2220
QY 2259 TGGGAGTGGCTCAGCTCTATTAAGACTTTTCTTAAATTTGAGAGCTTGAACCTTTGA 2318
DB 2221 TGGGAGTGGCTCAGCTCTATTAAGACTTTTCTTAAATTTGAGAGCTTGAACCTTTGA 2280
QY 2319 TATCCAGCTTAGCTGCTGCTCAGACTGAGCATGATCAAGAAAGACATGGGTTAAA 2378
DB 2281 TATCCAGCTTAGCTGCTGCTCAGACTGAGCATGATCAAGAAAGACATGGGTTAAA 2340
QY 2379 AGAACCAAGAGAGTTCGAAGAGCTATGCAACAGATCAAGAGCTTTAAAGACCAACA 2438
DB 2341 AGAACCAAGAGAGTTCGAAGAGCTATGCAACAGATCAAGAGCTTTAAAGACCAACA 2400
QY 2439 GAGTAAGGGACAGCTCTGGAGCCACGAGTCCAGAGTCTGCTGCTGCTGCTGCTGCTGCT 2498
DB 2401 GAGTAAGGGACAGCTCTGGAGCCACGAGTCCAGAGTCTGCTGCTGCTGCTGCTGCTGCT 2460
QY 2499 GAGGAAGATCTGCAACCTTGGGCTCCAGCGCATCTTCTACCTGAGCTTGAAGACTTGGT 2558
DB 2461 GAGGAAGATCTGCAACCTTGGGCTCCAGCGCATCTTCTACCTGAGCTTGAAGACTTGGT 2520
QY 2559 GTCTCCACTTCCATCATTTGAAGCTTCTCCTGGAACCCCTACCTTCTTAAATCAGGAGC 2618
DB 2521 GTCTCCACTTCCATCATTTGAAGCTTCTCCTGGAACCCCTACCTTCTTAAATCAGGAGC 2580
QY 2619 AGTGGAGGAGTGGCTGCTCCTCTCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2678
DB 2581 AGTGGAGGAGTGGCTGCTCCTCTCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
QY 2679 TTTGGAAACCTTATCATTTCTGCTTCTTAAAGAGGAAAAAGAGCTCCTGTAGAAAGCA 2738
DB 2641 TTTGGAAACCTTATCATTTCTGCTTCTTAAAGAGGAAAAAGAGCTCCTGTAGAAAGCA 2700
QY 2739 AAGACTTTCTTTTCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2798
DB 2701 AAGACTTTCTTTTCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2760
QY 2799 TATTGCTGTTGGGTTGTTGTTTATATTTAGGCAATTTGGGGATTTGGGTTGGGTTGGGTT 2858
DB 2761 TATTGCTGTTGGGTTGTTGTTTATATTTAGGCAATTTGGGGATTTGGGTTGGGTTGGGTT 2820
QY 2859 AGTTTCATGAGGGTTTCTTAAAGAAATGCTTAAAGAAAGCTTTTGGACAAATGCTATCCAG 2918
DB 2821 AGTTTCATGAGGGTTTCTTAAAGAAATGCTTAAAGAAAGCTTTTGGACAAATGCTATCCAG 2880
QY 2919 CAGGAAAAAAGAGGATATATATCTGTTTTTAAACTCTTTCTGCGGAAATCCAAATATAGT 2978


```
Db 2881 CAGGAAAAAAGGATATATATCTGTTTAAACTCTTCTGGGAATCCAAATTATAGT 2940
QY 2979 TGCCTTGATTTAAAAAAGCAAGCAAGCGGTGTTCCGCGAGGTAGGATGTCTCTTA 3038
Db 2941 TGCCTTGATTTAAAAAAGCAAGCAAGCGGTGTTCCGCGAGGTAGGATGTCTCTTA 3000
QY 3039 AAGATTGCTCCCTTCAAAATATGCTTCTGATCAAAAGTACGTATGGTCAACAG 3098
Db 3001 AAGATTGCTCCCTTCAAAATATGCTTCTGATCAAAAGTACGTATGGTCAACAG 3060
QY 3099 CGAGAAACTTCTCTTAAATTTCTTCTCTTCTTAAATTTAACTGTTGAAAGATGGAGG 3158
Db 3061 CGAGAAACTTCTCTTAAATTTCTTCTCTTAAATTTAACTGTTGAAAGATGGAGG 3120
QY 3159 ATTACTCAAAATCAGATGCAAAACAAATATGCTGTTTGCCTTCCATAAACAAGTGC 3218
Db 3121 ATTACTCAAAATCAGATGCAAAACAAATATGCTGTTTGCCTTCCATAAACAAGTGC 3180
QY 3219 AATTTTAAAGTCTGCTTACTAAGTCTGTTTAACTTCTCTTAACTCTCTTATATAGG 3278
Db 3181 AATTTTAAAGTCTGCTTACTAAGTCTGTTTAACTTCTCTTATATATAGG 3240
QY 3279 AATTTAAAGGAGGAGTCACTGTTAGCAATCAGACGTTTAAATATCCCTAGCAGGCTGT 3338
Db 3241 AATTTAAAGGAGGAGTCACTGTTAGCAATCAGACGTTTAAATATCCCTAGCAGGCTGT 3300
QY 3339 GTTCACTTCCCTGTCGATCCCTCTGAGGTATGCCCACATCAAGACTTTTAGGCCATTTC 3398
Db 3301 GTTCACTTCCCTGTCGATCCCTCTGAGGTATGCCCACATCAAGACTTTTAGGCCATTTC 3360
QY 3399 TTGATGGAACAGATCCCTGCGCTGACGTGTCAGCTATCTGAAAGTGGATCAGATTATA 3458
Db 3361 TTGATGGAACAGATCCCTGCGCTGACGTGTCAGCTATCTGAAAGTGGATCAGATTATA 3420
QY 3459 AACTCGATTACATGAATGTTTGTGTTGTTCTATCAACCCACAGAGTTCCCTAAA 3518
Db 3421 AACTCGATTACATGAATGTTTGTGTTGTTCTATCAACCCACAGAGTTCCCTAAA 3480
QY 3519 CTGCTTCAGTTATAGTAAGTCTGTTATATCAATTCAGAAAGCGCCATAAGTCAGTTGA 3578
Db 3481 CTGCTTCAGTTATAGTAAGTCTGTTATATCAATTCAGAAAGCGCCATAAGTCAGTTGA 3540
QY 3579 GTATTGATCCCTAGATAAGAAATCAATCAGCAGGAACTGTTATACAGGTTAAGCA 3638
Db 3541 GTATTGATCCCTAGATAAGAAATCAATCAGCAGGAACTGTTATACAGGTTAAGCA 3600
QY 3639 CCAGGCAATTAAGGATTTTATAGATATAATTTAAATTTTCTGTTA - TTGGTTAAGGAGA 3696
Db 3601 CCAGGCAATTAAGGATTTTATAGATATAATTTAAATTTTGGTAATTGGTTAAGGAGA 3660
QY 3697 -CAATTTTGGAGAGCAAGCAAA - - - TCTTTTAAAAAATAGTATGAATGTGAATACTAGA 3752
Db 3661 CCAATTTTGGAGAGCAAGCAATCTCTTTTAAAAAATAGTATGAATGTGAATACTAGA 3720
QY 3753 AAAGATTTAAAAAATAGTATGAATGTGAGTACTAGGAAGGAT 3794
Db 3721 AAAGATTTAAGAAATAGTATGAATGTGAGTACTAGGAAGGAT 3762
```

RESULT 8

```
ADF28866
ID ADF28866 standard; DNA; 4162 BP.
XX
AC ADF28866;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human nor-1 nuclear receptor nucleic acid 1651190.
XX
KW Human; nor-1; nuclear receptor; receptor; leukaemia; gene therapy;
KW cytosolic; haematopoietic cell; ds.
XX
```

```
OS Homo sapiens.
XX WO2003088812-A2.
XX 30-OCT-2003.
XX 15-APR-2003; 2003WO-US011804.
XX 17-APR-2002; 2002US-0373238P.
XX (BAYU ) BAYLOR COLLEGE MEDICINE.
XX (UNIW ) UNIV WASHINGTON.
XX Mullican SE, Conneely OM, Milbrandt J;
XX WPI; 2003-854017/79.
XX GENBANK; 1651190.
XX Inhibiting proliferation of a hematopoietic cell, useful for treating or
XX preventing leukemia, comprises modulating the level or activity of nor-1
XX and/or nur77 nuclear receptor.
XX Disclosure; SEQ ID NO 1; 101pp; English.
XX The present sequence is that of human nor-1 nuclear receptor nucleic acid
XX 1651190. Nor-1 has been identified as a molecular target for therapeutic
XX intervention in the treatment of myeloid leukaemia. The invention
XX provides methods of inhibiting the proliferation of a haematopoietic stem
XX cell or a haematopoietic myeloid cell by modulating the level of a nor-1
XX and/or nur77 nuclear receptor. This involves increasing the level of the
XX receptor polypeptide or polynucleotide, e.g. by administration of a
XX vector comprising the polynucleotide. A claimed method of treating
XX leukaemia comprises modulating a nor-1 and/or nur77 nuclear receptor in a
XX haematopoietic stem cell or myeloid cell. Also claimed are methods of
XX identifying an upregulator of expression of nor-1 and/or nur77, of
XX identifying a compound for the treatment of leukaemia, and of screening
XX for a compound for treatment of leukaemia, and a mouse model for
XX leukaemia comprising a mouse having defective nor-1 and nur77 nucleic
XX acid sequences.
XX Sequence 4162 BP; 1017 A; 1224 C; 991 G; 930 T; 0 U; 0 Other;
XX
XX Query Match 88.4%; Score 3352; DB 10; Length 4162;
XX Best Local Similarity 90.8%; Pred. No. 0;
XX Matches 3779; Conservative 0; Mismatches 15; Indels 369; Gaps 5;
QY 1 ATAAATGACGTGCCGAGAGAGCGGAAACGCGAGCGGAGAGCGGAGTCTCTGCTCT 60
Db 1 ATAAATGACGTGCCGAGAGAGCGGAAACGCGAGCGGAGAGCGGAGTCTCTGCTCT 60
QY 61 CCGCGCCCGCCACCCCTCCAGCTCTCTGCTCTCTCTCTCCGCTCCCATACACAGCGGCTCA 120
Db 61 CCGCGCCCGCCACCCCTCCAGCTCTCTGCTCTCTCTCTCCGCTCCCATACACAGCGGCTCA 120
QY 121 CACCGCTCTCCCTCACTCCGACACACACAGACACAAAGCGCGCACACAGGCTCCGACAC 178
Db 121 CACCGCTCTCCCTCACTCCGACACACACAGACACAAAGCGCGCACACAGGCTCCGACAC 180
QY 179 ACTTGCCTCTCCCGCGGCTCACACCCCTCTTGCCTGAGCCCTTGGCGGTGACGCGCG 238
Db 181 ACTTGCCTCTCCCGCGGCTCACACCCCTCTTGCCTGAGCCCTTGGCGGTGACGCGCG 240
QY 239 CGCGCAGCTGGAGCGCCCTCCCGGGCTCAGTTTGTCAACGCTGACGCGGTGCGGAGTGC 298
Db 241 CGCGCAGCTGGAGCGCCCTCCCGGGCTCAGTTTGTCAACGCTGACGCGGTGCGGAGTGC 300
QY 299 CGTGGAGTGGAAACAGCGCGGCACTCTCCCGCTGTGTCAAGCCCAAGCCAGGACGCC 358
Db 301 CGTGGAGTGGAAACAGCGCGGCACTCTCCCGCTGTGTCAAGCCCAAGCCAGGACGCC 360
QY 359 CGCGGAACCTCTCGGCTGTCTCTCCCATGAGTGGGATCGCAGCATCCCCCAGCGCG 418
Db 361 CGCGGAACCTCTCGGCTGTCTCTCCCATGAGTGGGATCGCAGCATCCCCCAGCGCG 420
```


QY 419 CTCACCGCTCCGGAGCGCTGGCTTGTATACCGCAGCCCTTCCGGGACAGAGCTGT 478
DB |||||
QY 421 CTCACCGCTCCGGAGCGCTGGCTTGTATACCGCAGCCCTTCCGGGACAGAGCTGT 480
DB |||||
QY 479 GACTCCCCCAGGTGCGATTTTGGGACAGCTCTCTAGAACTCGCTCTAAAGAGCGAAC 538
DB |||||
QY 481 GACTCCCCCAGGTGCGATTTTGGGACAGCTCTCTAGAACTCGCTCTAAAGAGCGAAC 540
DB |||||
QY 539 CGCACAGACTCAAGCCCACTGCGGAGAGGGCAGCCCGGCAAGCCCGGCGCTGAGC 598
DB |||||
QY 541 CGCCACAGACTCAAGCCCACTGCGGAGAGGGCAGCCCGGCAAGCCCGGCGCTGAGC 600
DB |||||
QY 599 CTGACCCCTTACCGGTGCGGCGACGACTGCGGCGCTTTCGCTCGCGGACGTCGCGTC 658
DB |||||
QY 601 CTGACCCCTTAGCGGTGCGGCGACGACTGCGGCGCTTTCGCTCGCGGACGTCGCGTC 660
DB |||||
QY 659 CTCCTACACTCTCAGCCTCGCTGAGAGACCCCGAGCCCAACATCAGCGCGCAAGAT 718
DB |||||
QY 661 CTCCTACACTCTCAGCCTCGCTGAGAGACCCCGAGCCCAACATCAGCGCGCAAGAT 720
DB |||||
QY 719 ACCCTCCAGATATGCCCTGCGTCCAAAGCCCAATATAGCCCTTCCGCTCCAGCTTCCAGTT 778
DB |||||
QY 721 ACCCTCCAGATATGCCCTGCGTCCAAAGCCCAATATAGCCCTTCCGCTCCAGCTTCCAGTT 780
DB |||||
QY 779 ATGCGGCGCAGACATACAGCTCGGAATACACACGAGATCATGAACCCCGACACCA 838
DB |||||
QY 781 ATGCGGCGCAGACATACAGCTCGGAATACACACGAGATCATGAACCCCGACACCA 840
DB |||||
QY 839 AGCTGACCATGACCTTGGAGCACTGAGATACGGCTACAGCCACAGCTCCCTGCCCA 898
DB |||||
QY 841 AGCTGACCATGACCTTGGAGCACTGAGATACGGCTACAGCCACAGCTCCCTGCCCA 900
DB |||||
QY 899 GCATCAGTACCTTCGTGGAGGCTACTCGAGCACTACGAACTCAAGCTTCTCCGCTGT 958
DB |||||
QY 901 GATCAGTACCTTCGTGGAGGCTACTCGAGCACTACGAACTCAAGCTTCTCCGCTGT 960
DB |||||
QY 959 ACCAAATGACGGCCCTTGATCAAAGTGGAGAGGGCGGGCGCCAGCTACCATACC 1018
DB |||||
QY 961 ACCAAATGACGGCCCTTGATCAAAGTGGAGAGGGCGGGCGCCAGCTACCATACC 1020
DB |||||
QY 1019 ATCAC 1078
DB |||||
QY 1021 ATCAC 1080
DB |||||
QY 1079 TTCCTCCAGCCTCCAGCCCGAGGACGAGGTGCTGCCAGCACCTCCATGACTTCAAGC 1138
DB |||||
QY 1081 TTCCTCCAGCCTCCAGCCCGAGGAGGAGAGGTGCTGCCAGCACCTCCATGACTTCAAGC 1140
DB |||||
QY 1139 AGTCCCCACCGTCCACCCGCCACACACCGCGCTTCCCGCGCAGCGCGGGGGCTTATGG 1198
DB |||||
QY 1141 AGTCCCCACCGTCCACCCGCCACACACCGCGCTTCCCGCGCAGCGCGGGGGCTTATGG 1200
DB |||||
QY 1199 ACGAGGACCTGCTCCGCGCCCGGCTGATCGCACCGCGCGCTGCTGGACCCGCCGA 1258
DB |||||
QY 1201 ACGAGGACCTGCTCCGCGCCCGGCTGATCGCACCGCGCGCTGCTGGACCCGCCGA 1260
DB |||||
QY 1259 TGAAGCGGTCCCGCGTGGCGCGCGCTTCCCGCTCTTCCACTTCAAGCCCTGCG 1318
DB |||||
QY 1261 TGAAGCGGTCCCGCGTGGCGCGCGCTTCCCGCTCTTCCACTTCAAGCCCTGCG 1320
DB |||||
QY 1319 CGCCGATCCCCCGCGCCAGCCCGCGCGCGGCGACCACTCGGCTACGACCGGAGCG 1378
DB |||||
QY 1321 CGCCGATCCCCCGCGCCAGCCCGCGCGCGGCGACCACTCGGCTACGACCGGAGCG 1380
DB |||||
QY 1379 CGCTGCGCGCTCAGCTGCGCTGGAGCGCGACCGCGCGCGGCGACCGGCGCGCG 1438
DB |||||
QY 1381 CGCTGCGCGCTCAGCTGCGCTGGAGCGCGACCGCGCGCGGCGACCGGCGCGCG 1440
DB |||||
QY 1439 CGCTTGAGACCAACCGTACGGCTGCGCTGCGCAAGAGGCGCGCGCTGCGCTTCC 1498
DB |||||
QY 1441 CGCTTGAGGCGCACCGTACGGCTGCGCTGCGCAAGAGGCGCGCGCTGCGCTTCC 1500
DB |||||

QY 1499 GGCCTCTCGGCTCACGCGCTCCCTACCGGCTCCAGCTGCTGGGCGAGAGTCCAGCC 1558
DB |||||
QY 1501 GGCCTCTCGGCTCACGCGCTCCCTACCGGCTCCAGCTGCTGGGCGAGAGTCCAGCC 1560
DB |||||
QY 1559 TGCCGTGCGCGCCAGCAGAGAGCTGCTGCTGCGGAGGACAGCTGTCGCTGTGCGGG 1618
DB |||||
QY 1561 TGCCGTGCGCGCCAGCAGAGAGCTGCTGCTGCGGAGGACAGCTGTCGCTGTGCGGG 1620
DB |||||
QY 1619 ACAACGCGCGCTGCCAGCATTACGCGCTGCGAACTCGGAGGGTGAAGGGCTTTTCA 1678
DB |||||
QY 1621 ACAACGCGCGCTGCCAGCATTACGCGCTGCGAACTCGGAGGGTGAAGGGCTTTTCA 1680
DB |||||
QY 1679 AGAGAAAGTGCAGAAATAATGCTTTCCTGGGCAATAAATAATGCTTTCAGTAG 1738
DB |||||
QY 1681 AGAGAAAGTGCAGAAATAATGCTTTCCTGGGCAATAAATAATGCTTTCAGTAG 1740
DB |||||
QY 1739 ACAAGAGCGTGCAGAAACCGATGTCAGTACTGTCGATTTTCAAGTGTCTCAGTGTGAA 1798
DB |||||
QY 1741 ACAAGAGCGTGCAGAAACCGATGTCAGTACTGTCGATTTTCAAGTGTCTCAGTGTGAA 1800
DB |||||
QY 1799 TGGTAAAGAGTGTTCGCTACAGATAGTCTGAAGGAGGAGAGTGTCTGCTTCCA 1858
DB |||||
QY 1801 TGGTAAAGAGTGTTCGCTACAGATAGTCTGAAGGAGGAGAGTGTCTGCTTCCA 1860
DB |||||
QY 1859 AACCAAGAGCCATTTACACAGGAACTTCTCAGCCCTCTCCACCTTCTCCTCAATCT 1918
DB |||||
QY 1861 AACCAAGAGCCATTTACACAGGAACTTCTCAGCCCTCTCCACCTTCTCCTCAATCT 1920
DB |||||
QY 1919 GCATGATGAATGCGCTTGTGCGAGCTTAAACAGACTCAACACCCAGAGATCTTGATTT 1978
DB |||||
QY 1921 GCATGATGAATGCTTGTGCGAGCTTAAACAGACTCAACACCCAGAGATCTTGATTT 1980
DB |||||
QY 1979 CCAGATAGTGTCCCACTGACAGGCTGTGCGACACAGATGTGAGCATGTGCAACAAT 2038
DB |||||
QY 1981 CCAGATAGTGTCCCACTGACAGGCTGTGCGACACAGATGTGAGCATGTGCAACAAT 2040
DB |||||
QY 2039 TCTCAACCTCTGACAGCCTCAATTG----- 2065
DB |||||
QY 2041 TCTCAACCTCTGACAGCCTCAATTGATGATCCAGAAGCTGGGCAAGAGGATTCGG 2100
DB |||||
QY 2066 ----- 2065
DB |||||
QY 2101 GATTTACTGATCTCCCAAGAGATCAGACATTAATTATTTGATCAAGCTTTTGGAGC 2160
DB |||||
QY 2066 ----- 2065
DB |||||
QY 2161 TGTTTGTCTCAGACTTTCCATCAGTCAAACTGCTGGAAGATAAGTTTGTGTTCTGCA 2220
DB |||||
QY 2066 ----- 2065
DB |||||
QY 2221 ATGGAATTTGCTGCACTGACTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCA 2280
DB |||||
QY 2066 ----- 2065
DB |||||
QY 2281 TTAAGACTTTTCTCTAAATTTGACAGCCTGAACCTTGATATCCAAGCCTTAGCCTGCC 2340
DB |||||
QY 2066 ----- 2065
DB |||||
QY 2341 TGTACAGACTGAGCATGATCAGAGAGACATGGGTTAAAGAAACCAAGAGAGTTCGAAG 2400
DB |||||
QY 2066 -----ATGATATCCAGAAGCTGGGCAAGAGATTTCCGG 2098
DB |||||
QY 2401 TCTACAACTCTGACAGCCTCCATTGATGATCCAGAAGCTGGGCAAGAGATTCGG 2460
DB |||||
QY 2099 GATTTACTGATCTCCCAAGAGATCAGACATTAATTATTTGAATCAGCCTTTTGGAGC 2158
DB |||||
QY 2461 GATTTACTGATCTCCCAAGAGATCAGACATTAATTATTTGAATCAGCCTTTTGGAGC 2520
DB |||||
QY 2159 TGTTTGTCTCAGACTTTTCCATCAGGTCAAACTGCTGGAAGATAAGTTTGTGTTCTGCA 2218
DB |||||
QY 2521 TGTTTGTCTCAGACTTTTCCATCAGTCAAACTGCTGGAAGATAAGTTTGTGTTCTGCA 2580
DB |||||
QY 2219 ATGGAATTTGCTGCACTTCAATGCTTTCGTGGATTTGGGAGTGGCTCGACTCTA 2278
DB |||||

```
Db 2581 ATGGACTTGTCTGCATCGACTTCAGTGCCTTCGTGGATTTGGGGAGTGGCTCGACTCTA 2640
QY 2279 TTAAGAATTTTCTTAAATTTGCAGAGCTGAACCTTGATATCCAAGCTTAGCCTGCC 2338
Db 2641 TTAAGACTTTTCTTAAATTTGCAGAGCTGAACCTTGATATCCAAGCTTAGCCTGCC 2700
QY 2339 TGTGACACTGAGCATGATCAACAGAAAGACATGGGTTAAAGAACCAAAAGAGAGTCGAAG 2398
Db 2701 TGTGACACTGAGCATGATCAACAGAAAGACATGGGTTAAAGAACCAAAAGAGAGTCGAAG 2760
QY 2399 AGCTATGCAACAAGATCAACAAGCAGTTTAAAGACCAACAGAGTAAGGGACAGGCTCTGG 2458
Db 2761 AGCTATGCAACAAGATCAACAAGCAGTTTAAAGACCAACAGAGTAAGGGACAGGCTCTGG 2820
QY 2459 AGCCCAACGAGTCCAAGGTCCTGGGTCCTGGTGAACCTGAGGAGATCTGCACCCCTGG 2518
Db 2821 AACCACCAAGTCCAAGGTCCTGGTGGTCCCTGGTGAACCTGAGGAGATCTGCACCCCTGG 2880
QY 2519 GCCTCAGCGCATCTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCGATCATTTG 2578
Db 2881 GCCTCAGCGCATCTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCGATCATTTG 2940
QY 2579 ACAAGCTCTTCTGGACACCTTACCTTCTTAATCAGGACAGTGGAGCAGTGAAGTGCCT 2638
Db 2941 ACAAGCTCTTCTGGACACCTTACCTTCTTAATCAGGACAGTGGAGCAGTGAAGTGCCT 3000
QY 2639 CCTCTCCTAGCACCTCTTGTCTACGACGAAAGGATAGGTTTGGAAACCTATCATTTCC 2698
Db 3001 CCTCTCCTAGCACCTCTTGTCTACGACGAAAGGATAGGTTTGGAAACCTATCATTTCC 3060
QY 2699 TGTCTCTCTTAAGAGGAAAGAGCCTCTCTGTAGAAAGCAAGACTTCTTTTTTTCTG 2758
Db 3061 TGTCTCTCTTAAGAGGAAAGAGCCTCTCTGTAGAAAGCAAGACTTCTTTTTTTCTG 3120
QY 2759 GCTCTTTCTCTTACACCTTAAGACGAGAAATCTGCAGAGTATGTGTGGGTTGTCT 2818
Db 3121 GCTCTTTCTCTTACACCTTAAGACGAGAAATCTGCAGAGTATGTGTGGGTTGTCT 3180
QY 2819 TTATATTAGGCAATTTGGGGATGGGGTGGAGGGGTATAGTTTATGAGGGTCTTCTAA 2878
Db 3181 TTATATTAGGCAATTTGGGGATGGGGTGGAGGGGTATAGTTTATGAGGGTCTTCTAA 3240
QY 2879 GAAATTGTCAACAAGCACTTTTGACAAATGTATCCACGACGAGGAAAAAGGATAATA 2938
Db 3241 GAAATTGTCAACAAGCACTTTTGACAAATGTATCCACGACGAGGAAAAAGGATAATA 3300
QY 2939 TAACCTGTTTAAACCTTTCTGGGGAATCCAAATATAGTTGCTTTGTTTAAACAA 2998
Db 3301 TAACCTGTTTAAACCTTTCTGGGGAATCCAAATATAGTTGCTTTGTTTAAACAA 3360
QY 2999 GAAACAGCAAGGTTGTTCCGACAGGTAGGATGTGTCTTAAAGATTTGGTCCCTTGAAT 3058
Db 3361 GAAACAGCAAGGTTGTTCCGACAGGTAGGATGTGTCTTAAAGATTTGGTCCCTTGAAT 3420
QY 3059 ATGCTTCTGTATCAAGGATCGTATGTGGTGCAACCAAGGACAGAACTTCTTTTAAT 3118
Db 3421 ATGCTTCTGTATCAAGGATCGTATGTGGTGCAACCAAGGACAGAACTTCTTTTAAT 3480
QY 3119 TCCTTCTCTCTTATTTTAAACAAATGGTGAAGATGGAGATTAACCTACAAATCAGACAT 3178
Db 3481 TCCTTCTCTCTTATTTTAAACAAATGGTGAAGATGGAGATTAACCTACAAATCAGACAT 3540
QY 3179 GCGAAACAAATATGCTGTTTCTTCCATAACCAAGTGCATTTTAAAGTGTGCT 3238
Db 3541 GCGAAACAAATATGCTGTTTCTTCCATAACCAAGTGCATTTTAAAGTGTGCT 3600
QY 3239 TACTAAGTCTGTTTATTAACCTCTCTTTTATTTCTATATGGAATATAAGAGGAGTCA 3298
Db 3601 TACTAAGTCTGTTTATTAACCTCTCTTTTATTTCTATATGGAATATAAGAGGAGTCA 3660
QY 3299 TGTTAGCAATGACAGCTTAATATCCCTAGCAGAGCTGTGTTACCTTCCCTGTGATC 3358
```

```
Db 3661 TGTTAGCAATGACAGCTTAATATCCCTAGCAGAGCTGTGTTACCTTCCCTGTGATC 3720
QY 3359 CCTTCTGAGGTATGCCCATCCAAGACTTTTATGGCCATTCTTATGAGAACCGATCCCTG 3418
Db 3721 CCTTCTGAGGTATGCCCATCCAAGACTTTTATGGCCATTCTTATGAGAACCGATCCCTG 3780
QY 3419 CCTGACTGTCCAGCTATCCTGAAAGTGGATCAGATTATAAACTGGATTACATGTAAC 3478
Db 3781 CCTGACTGTCCAGCTATCCTGAAAGTGGATCAGATTATAAACTGGATTACATGTAAC 3840
QY 3479 TTTTGGTGTGTTCTTATCAACCCACCCAGAGTTCCCTAAACTTGCCTTCACTTATAGTAAC 3538
Db 3841 TTTTGGTGTGTTCTTATCAACCCACCCAGAGTTCCCTAAACTTGCCTTCACTTATAGTAAC 3900
QY 3539 TGACTGGTATATTCATTTCAGAAAGCCCATTAAGTCAAGTTCAGTATTTGATCCCTAGATAAG 3598
Db 3901 TGACTGGTATATTCATTTCAGAAAGCCCATTAAGTCAAGTTCAGTATTTGATCCCTAGATAAG 3960
QY 3599 AACATGCAAAATCAGCAGGAACTGGTTCATACAGGGTAAGCACCAGGACAAATAGGATTTT 3658
Db 3961 AACATGCAAAATCAGCAGGAACTGGTTCATACAGGGTAAGCACCAGGACAAATAGGATTTT 4020
QY 3659 TATAGATATAATTTAAATTTTCTTA--TTGGTTAAGGAGA-CAATTTTGGAGGCAAGCA 3715
Db 4021 TATAGATATAATTTAAATTTTCTTA--TTGGTTAAGGAGA-CAATTTTGGAGGCAAGCA 4080
QY 3716 AA---TCTTTTAAAAAATAGTATGAATGTGAATACTAGAAAAGATTTTAAAAAATAGTAT 3772
Db 4081 AATCTCTTTTAAAAAATAGTATGAATGTGAATACTAGAAAAGATTTTAAAAAATAGTAT 4140
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794
Db 4141 GAGTGTGAGTACTAGGAAGGAT 4162
```

RESULT 9

ABU61892

ID ABL61892 standard; DNA; 4977 BP.

XX

AC ABL61892;

XX

DT 15-MAY-2002 (first entry)

XX

DE Colon adenocarcinoma related gene sequence SEQ ID NO:229.

XX

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

KW gene; ds.

XX

OS Homo sapiens.

XX

PN WO200194629-A2.

XX

PD 13-DEC-2001.

XX

XX 30-MAY-2001; 2001WO-US010838.

PR

PR 05-JUN-2000; 2000US-0209473P.

PR

PR 18-SEP-2000; 2000US-0209531P.

PR

PR 18-SEP-2000; 2000US-0233133P.

PR

PR 20-SEP-2000; 2000US-0233617P.

PR

PR 20-SEP-2000; 2000US-0234009P.

PR

PR 20-SEP-2000; 2000US-0234034P.

PR

PR 20-SEP-2000; 2000US-0234052P.

PR

PR 22-SEP-2000; 2000US-0234505P.

PR

PR 22-SEP-2000; 2000US-0234567P.

PR

PR 25-SEP-2000; 2000US-0234923P.

PR

PR 25-SEP-2000; 2000US-0234942P.

PR

PR 25-SEP-2000; 2000US-0235077P.

PR

PR 25-SEP-2000; 2000US-0235082P.

PR

PR 25-SEP-2000; 2000US-0235134P.

PR

PR 25-SEP-2000; 2000US-0235280P.

PR

PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 03-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 229; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 4977 BP; 1394 A; 1147 C; 1042 G; 1394 T; 0 U; 0 Other;

Query Match 75.8%; Score 2875; DB 6; Length 4977;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;

QY 733 CCCTGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGG--CGCAGA 790
DB 94 CCTGTCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGTCGTCTCAGA 153
QY 791 CATACAGCTCGGAATACACCGAGATCATGAACCCGAGCTACACCAAGCTGACCATGG 850
DB 154 CATACAGCTCGGAATACACCGAGATCATGAACCCGAGCTACACCAAGCTGACCATGG 213

QY 851 ACCTTGGCAGCACTGAGATCAACGGCTACAGCCACACAGTCCCTGCCCCAGCATCAGTACCT 910
DB 214 ACCTTGGCAGCACTGAGATCAACGGCTACAGCCACACAGTCCCTGCCCCAGCATCAGTACCT 273
QY 911 TCCTGGAGGGCTACTCGAGCAACTACGAACCTCAAGCCTTCTCGGTGTACCAAAATGCAGC 970
DB 274 TTGTGGAGGGCTACTCGAGCAACTACGAACCTCAAGCCTTCTCGGTGTACCAAAATGCAGC 333
QY 971 GGCCCTTTGATCAAAAGTGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1030
DB 334 GGCCCTTTGATCAAAAGTGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 393
QY 1031 ACACACACACACACACACATCACAGCAGCAGCATCAGCAGCAGCATCCTATTCCTCCAGCCT 1090
DB 394 ACCACACACACACACACATCACAGCAGCAGCATCAGCAGCAGCATCCTATTCCTCCAGCCT 453
QY 1091 CCAGCCGGAGGAGAGGTGTGCGCCAGCAGCCTCCATGTACTTCAAGCAGTCCCCACCGT 1150
DB 454 CCAGCCGGAGGAGAGGTGTGCGCCAGCAGCCTCCATGTACTTCAAGCAGTCCCCACCGT 513
QY 1151 CCAGCCGGAGGAGAGGTGTGCGCCAGCAGCCTCCATGTACTTCAAGCAGTCCCCACCGT 1210
DB 514 CCAGCCGGAGGAGAGGTGTGCGCCAGCAGCCTCCATGTACTTCAAGCAGTCCCCACCGT 573
QY 1211 CCTCGGGCCCGGCTGATGCAACCGGCCCGGCTGTGTGGAACCGCCGATGAAGCGGTCC 1270
DB 574 CCTCGGGCCCGGCTGATGCAACCGGCCCGGCTGTGTGGAACCGCCGATGAAGCGGTCC 633
QY 1271 CCAGGTGGCGGGCGGGCTTCCGGCTTCTTCCACTTCAAGCCCTCGCGCGCGGATCCCC 1330
DB 634 CCAGGTGGCGGGCGGGCTTCCGGCTTCTTCCACTTCAAGCCCTCGCGCGCGGATCCCC 693
QY 1331 CCGGCCCGCAGCCCGGGCGGGCCACACCTTCGGCTACGACCCGACCGCCCGCTGCGGCG 1390
DB 694 CCGGCCCGCAGCCCGGGCGGGCCACACCTTCGGCTACGACCCGACCGCCCGCTGCGGCG 753
QY 1391 TCAGCTTGCCTGGAGCGCGCAGCGCGCGGGGAGCGCAGCGCGCGCTTGAAGACC 1450
DB 754 TCAGCTTGCCTGGAGCGCGCAGCGCGCGGGGAGCGCAGCGCGCGCTTGAAGGCC 813
QY 1451 ACCGTAGCGGCTCGCTGCGCCAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1510
DB 814 ACCGTAGCGGCTCGCTGCGCCAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 873
QY 1511 TCAGCCCTCCCTTACCGCGCTCCAGCTGTGGGGGAGAGTCCAGCAGCTGCGGTGCGGCG 1570
DB 874 TCAGCCCTCCCTTACCGCGCTCCAGCTGTGGGGGAGAGTCCAGCAGCTGCGGTGCGGCG 933
QY 1571 CCAGCAGAGCTCGCTGCTGCGCAGGCGCAGTGTGCGGTGCGGGGAGCAGCGCGCT 1630
DB 934 CCAGCAGAGCTCGCTGCTGCGCAGGCGCAGTGTGCGGTGCGGGGAGCAGCGCGCT 993
QY 1631 GCCAGCATCAGCGCTGCGAACCTTCGAGGGCTTCAAGGGCTTTTTCAGAGAGACAGTGC 1690
DB 994 GCCAGCATCAGCGCTGCGAACCTTCGAGGGCTTCAAGGGCTTTTTCAGAGAGACAGTGC 1053
QY 1691 AGAAAAATGCAAAAATATGTTTTCCTGGCAATAAAAAATGCCCCAGTAGACAGACGTC 1750
DB 1054 AGAAAAATGCAAAAATATGTTTTCCTGGCAATAAAAAATGCCCCAGTAGACAGACGTC 1113
QY 1751 GAAACCGATGTGCTAGTGTGATTTCAAGAGTGTCTCAGTGTGGAATGTTAAAGAG 1810
DB 1114 GAAACCGATGTGCTAGTGTGATTTCAAGAGTGTCTCAGTGTGGAATGTTAAAGAG 1173
QY 1811 TTGTCCGTACAGATAGTCTGAAAGGGAGGAGGTGCTGCTGCTTCCAAAACCAAGAGCC 1870
DB 1174 TTGTCCGTACAGATAGTCTGAAAGGGAGGAGAGGTGCTGCTGCTTCCAAAACCAAGAGCC 1233
QY 1871 CATTACAAAGGAACCTTCTCAGCCCTCTCAGCTTCTCCTTCCAAATTCGATGATGAATG 1930
DB 1234 CATTACAAAGGAACCTTCTCAGCCCTCTCAGCCCTCTCCTTCCAAATTCGATGATGAATG 1293
QY 1931 CCCTGTGCGGAGCTTTTAAACAGACTCAACCCAGAGATCTTGAATTTATTCAGATACTGTC 1990

[illegible]

Db	2374	GTGTTCCGACGGGTAGGA	TGTGTTTAAAGATTGGTCCCTCGAAATATATGCTTCCTGTA	2433
Qy	3071	TCAAAGTAGTATGTTGGTCAAA	CAAGGCAGAAACTTCCTTTAAATTTCTTCTTCCTT	3130
Db	2434	TCAAAGGTACGTATGTTGGTCAAA	CAAGGCAGAACTTCTTTAATTTCTTCTTCCTT	2493
Qy	3131	TATTTTAACAAAATGGTGAAGATGGAGGATTA	CTTCAAAATCAGACATGATGCAAAAAA	3190
Db	2494	TATTTTAACAAAATGGTGAAGATGGAGGATTA	CTTCAAAATCAGACATGATGCAAAAAA	2551
Qy	3191	ATGGCTGTTTGGCTTCCATAACAAAGTGCATTTT	TAAAGTGCCTTCTTACTAAGTCTTG	3250
Db	2552	ATGGCTGTTTGGCTTCCATAACAAAGTGCATTTT	TAAAGTGCCTTCTTACTAAGTCTTG	2608
Qy	3251	TTTATTAACTCTCTCTTTATTTCTATATAGGAAT	AAAAAGGAGGCAGTCATGTTAGCAAAATG	3310
Db	2609	T--TATCTCTTATCTATATGCCGAATTA	AAAAAGGAGGCAGTCATGTTAGCAAAATG	2665
Qy	3311	ACAGTTTAATATCCCTAGCAGAGGCTGTGTCA	CTTCCCTGTCGATCCCTCTGAGGTA	3370
Db	2666	ACAGTTTAATATCCCTAGCAGAGGCTGTGTCA	CTTCCCTGTCGATCCCTCTGAGGTA	2725
Qy	3371	TGGCCCATCAAGACTTTTAGGCGCATCTTGAT	TGGAACAGATCCCTGCCCTGACTGTCC	3430
Db	2726	TGGCCCATCAAGACTTTTAGGCGCATCTTGAT	TGGAACAGATCCCTGCCCTGACTGTCC	2785
Qy	3431	AGCTATCCTCAAAAGTGGATCAGATTAATAAA	CTGGATTACATGTAACCTGTTTGGTTGTGT	3490
Db	2786	AGCTATCCTCAAAAGTGGATCAGATTAATAAA	CTGGATTACATGTAACCTGTTTGGTTGTGT	2845
Qy	3491	TCTATCAACCCCAACAGAGTTCCTTAACT	TGCTTTCAGTTATAGTAACTGGTATAT	3550
Db	2846	TCTATCAACCCCAACAGAGTTCCTTAACT	TGCTTTCAGTTATAGTAACTGGTATAT	2904
Qy	3551	TCATTGAGAGCGGCATAGTCACTTGGATTTTG	ATCCCTAGATAAGACATGCAATC	3610
Db	2905	TCATTGAGAGCGGCATAGTCACTTGGATTTTG	ATCCCTAGATAAGACATGCAATC	2959
Qy	3611	AGCAGGAATCTGGTCATACAGGGTAAGCACC	AGGGACAATAAGGATTTTTATAGATATAAT	3670
Db	2960	TCAGAGACTGGTCATACAGGGTAAGCACC	AGGGACAATAAGGATTTTTATAGATATAAT	3019
Qy	3671	TTAATTTTCTTATTTGGTTAAGGAGACAA	TTTTTGGAGGACAGCAAA---TCTTTTAAA	3727
Db	3020	TTAATTTTCTTATTTGGTTAAGGAGACAA	TTTTTGGAGGACAGCAAACTCTCTTTTAAA	3079
Qy	3728	AAATAGTATGAATGTGAATCTAGAAAGATTT	AAAAAATAGTATGAGTGTGAGTACTAG	3787
Db	3080	AAATAGTATGAATGTGAATCTAGAAAGATTT	AAAAAATAGTATGAGTGTGAGTACTAG	3139
Qy	3788	GAAGGAT	3794	
Db	3140	GAAGGAT	3146	
RESULT 10				
ABK84229				
ID	ABK84229 standard; cDNA; 4977 BP.			
XX	ABK84229;			
XX	14-AUG-2002 (first entry)			
XX	Human cDNA differentially expressed in granulocytic cells #800.			
XX	Human; ss; granulocytic cell; DNA chip; bacterial infection;			
KW	viral infection; parasitic infection; protozoal infection;			
KW	fungal infection; sterile inflammatory disease; psoriasis;			
KW	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;			
KW	cardiac reperfusion injury; renal reperfusion injury; ARDS;			
KW	adult respiratory distress syndrome; inflammatory bowel disease;			
KW	Crohn's disease; ulcerative colitis; periodontal disease;			
KW	granulocyte activation; chronic inflammation; allergy.			

XX	Homo sapiens.	QY	851	ACCTTGGCAGCCTGAGATCAACGGCTACAGCCACACAGTCCCTGCTCCAGCAGATCAGTACCT	910
OS		Db	214	ACCTTGGCAGCCTGAGATCAACGGCTACAGCCACACAGTCCCTGCTCCAGCAGATCAGTACCT	273
PN	WO200228999-A2.	QY	911	TCGTGGAGGGCTACTCGAGCAACTACGAACTCAGAGCCTTCTGCTGGTGTACCAATGACAGC	970
XX		Db	274	TTGTGGAGGGCTACTCGAGCAACTACGAACTCAGAGCCTTCTGCTGGTGTACCAATGACAGC	333
PD	11-APR-2002.	QY	971	GGCCCTTTGATCAAAAGTGGAGGAGGGGGGGCCAGCTACCATCACCATCAACACACACC	1030
XX		Db	334	GGCCCTTTGATCAAAAGTGGAGGAGGGGGGGCCAGCTACCATCACCATCAACACACACC	393
PF	03-OCT-2001; 2001WO-US030821.	QY	1031	ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	1090
XX		Db	394	ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	453
PA	(GENE-) GENE LOGIC INC.	QY	1091	CCAGCCCGAGGAGCAGAGTGTCTGCCAGCAGCCTCCATGTACTTCAAGCAGTCCCAACCGT	1150
XX		Db	454	CCAGCCCGAGGAGCAGAGTGTCTGCCAGCAGCCTCCATGTACTTCAAGCAGTCCCAACCGT	513
PI	Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;	QY	1151	CCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	1210
XX		Db	514	CCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	573
DR	WPI; 2002-435328/46.	QY	1211	CCTCGGCGCCCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1270
XX		Db	574	CCTCGGCGCCCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	633
PT	Detecting granulocyte activation by detecting differential expression of	QY	1271	CCAGGTTGGCGCGCGCGGTTCCTCCAGCTTCCAGCTTCAAGCCCTCCGCGCGCGCATCCCC	1330
XX	genes associated with granulocyte activation, which serves as diagnostic	Db	634	CCAGGTTGGCGCGCGCGGTTCCTCCAGCTTCCAGCTTCAAGCCCTCCGCGCGCGCATCCCC	693
PT	markers that is useful for monitoring disease states and drug toxicity.	QY	1331	CCGCGCCCGAGCCCG	1390
XX		Db	694	CCGCGCCCGAGCCCG	753
PS	Claim 1; SEQ ID NO 800; 114pp; English.	QY	1391	TCAGCCTTCCGCTCGGAGCCG	1450
XX		Db	754	TCAGCCTTCCGCTCGGAGCCG	813
CC	The invention relates to detecting (M1) granulocyte (GC) activation	QY	1451	ACCGTACGGGCTGCGGCTGCGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1510
CC	(GCA), by detecting the level of expression of gene(s) (Gs) identified by	Db	814	ACCGTACGGGCTGCGGCTGCGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	873
CC	DNA chip analysis as given in the specification, and comparing the	QY	1511	TCAGCCTTCCGCTCGGAGCCG	1570
CC	expression level to an expression level in an unactivated GC, where	Db	874	TCAGCCTTCCGCTCGGAGCCG	933
CC	differential expression of Gs is indicative of GCA. Also included are	QY	1571	CCAGCAGGAGTCTGCTGCTGCGGAGGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1630
CC	modulating (M2) GA by contacting GC with an agent that alters the	Db	934	CCAGCAGGAGTCTGCTGCTGCGGAGGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	993
CC	expression of at least one gene in Gs; (2) screening (M3) for an agent	QY	1631	CCAGCAGTACGGGCTGCGGACCTCGGAGGGGCTGCAAGGGGCTTTTCAAGAGAACAGTGC	1690
CC	capable of modulating GCA or an inflammation (especially chronic) in a	Db	994	CCAGCAGTACGGGCTGCGGACCTCGGAGGGGCTGCAAGGGGCTTTTCAAGAGAACAGTGC	1053
CC	tissue, an allergic response in a subject, exposure of a subject to a	QY	1691	AGAAAATGCAAAAATATGTTTTCCTGCGCAAAATAAAAAATGCGCCAGTACAGAGACGCTC	1750
CC	pathogen or sterile inflammatory disease using the gene expression	Db	1054	AGAAAATGCAAAAATATGTTTTCCTGCGCAAAATAAAAAATGCGCCAGTACAGAGACGCTC	1113
CC	profile; (3) detecting (M4) an inflammation (especially chronic) in a	QY	1751	GAACCCGATGTCAGTACGATGTCGATTTTTCAGAGTGTCTCAGTGTGGTAAAGAGAG	1810
CC	tissue, an allergic response in a subject, exposure of a subject to a	Db	1114	GAACCCGATGTCAGTACGATGTCGATTTTTCAGAGTGTCTCAGTGTGGTAAAGAGAG	1173
CC	pathogen or sterile inflammatory disease, by detecting the level of	QY	1811	TTGTCCGTACAGATGTCGATGTCGATTTTTCAGAGTGTCTCAGTGTGGTAAAGAGAG	1870
CC	expression in a sample of the tissue of gene(s) from Gs, where the level	Db	1174	TTGTCCGTACAGATGTCGATGTCGATTTTTCAGAGTGTCTCAGTGTGGTAAAGAGAG	1233
CC	of expression of the gene is indicative of inflammation; (4) treating	QY	1871	CATTACAAACGAGAACCTTCTCAGCCCTCTCCAACCTTCTCCTTCCCAATCTGCATGATG	1930
CC	(M5) an inflammation (especially chronic) or in a tissue, an allergic	Db	1234	CATTACAAACGAGAACCTTCTCAGCCCTCTCCAACCTTCTCCTTCCCAATCTGCATGATG	1293
CC	response in a subject, exposure of a subject to a pathogen or sterile	QY	1931	CCCTTGTCCGAGCTTTTAAACAGACTCAACCCAGAGATCTTGAATTAATTCAGATACTGTC	1990
CC	inflammatory disease, by contacting a tissue having inflammation with an				
CC	agent that modulates the expression of gene(s) from Gs in the tissue. M1				
CC	is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful				
CC	for screening an agent capable of modulating GCA preferably in an				
CC	inflammation in a tissue; M4 is useful for detecting an inflammation				
CC	(especially chronic) in a tissue, an allergic response in a subject,				
CC	exposure of a subject to a pathogen or sterile inflammatory disease (e.g.				
CC	psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,				
CC	cardiac reperfusion injury, renal reperfusion injury, ARDS, adult				
CC	respiratory distress syndrome, inflammatory bowel disease, Crohn's				
CC	disease, ulcerative colitis, periodontal disease; also bacterial				
CC	infection, viral infection, parasitic infection, protozoal infection,				
CC	fungal infection and M5 is useful for treating one of the above				
CC	conditions. The present sequence represents a gene differentially				
CC	expressed in granulocytes. Note: The sequence data for this patent did				
CC	not form part of the printed specification, but was obtained in				
CC	electronic format directly from WIPO at				
CC	ftp.wipo.int/pub/published_pct_sequences				
XX		XX			
SQ	Sequence 4977 BP; 1394 A; 1147 C; 1042 G; 1394 T; 0 U; 0 Other;				
	Query Match 75.8%; Score 2875; DB 6; Length 4977;				
	Best Local Similarity 98.1%; Pred. No. 0;				
	Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;				
QY	733 CCCTCGCTCGAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGG--CGCAGA				790
Db					
QY	94 CCGTGTCTCGAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGTCGTCAGA				153
Db					
QY	791 CATACAGCTCGGAATACACCGAGAGATCATGAACCCGAGCTACACCAAGCTGACCATGG				850
Db					
Db	154 CATACAGCTCGGAATACACCGAGAGATCATGAACCCGAGCTACACCAAGCTGACCATGG				213

XX PD 25-MAR-2004.

XX PF 15-SEP-2003; 2003WO-US029097.

XX PR 16-SEP-2002; 2002US-0411392P.

XX PA (GETH) GENENTECH INC.

XX PI Chiu H, Clark H, Dennis K, Fong S, Schoenfeld JR, Wood WI, Wu TD;

XX PI WPI; 2004-329389/30.

XX DR P-PSDB; ADL83262.

XX PT New PRO polypeptide, useful for diagnosing and treating a B cell related disorder, e.g. Burkitt's lymphoma, rheumatoid arthritis, autoimmune mediated hemolytic anemia, myasthenia gravis or ankylosing spondylitis.

XX PS Claim 2; Fig 463; 695pp; English.

XX CC The present invention relates to PRO proteins and their coding sequences. The PRO proteins are useful for diagnosing and treating a B cell related disorder, e.g. X-linked infantile hypogammaglobulinemia, polysaccharide antigen unresponsiveness, selective IgA deficiency, selective IgM deficiency, selective deficiency of IgG subclasses, immunodeficiency with hyper IgM, transient hypogammaglobulinemia of infancy, Burkitt's lymphoma, intermediate lymphoma, follicular lymphoma, type II hypersensitivity, rheumatoid arthritis, autoimmune mediated haemolytic anaemia, myasthenia gravis, hypoadrenocorticism, glomerulonephritis, or ankylosing spondylitis. The PRO proteins are also useful for preparing a medicament for treating a condition that is responsive to the PRO protein, e.g. cancer or immune-mediated inflammatory diseases. The PRO coding sequences are useful as hybridization probes in chromosome and gene mapping, in preparing PRO proteins, or in generating transgenic animals or knockout animals, which in turn are useful in the development and screening of therapeutically useful reagents.

XX SQ Sequence 4977 BP; 1394 A; 1147 C; 1042 G; 1394 T; 0 U; 0 Other;

Query Match 75.8%; Score 2875; DB 12; Length 4977;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;

QY 733 CCCTGCGTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGG--CGCAGA 790

DB 94 CCTGTCGTCCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGTGTGCTAGA 153

QY 791 CATACAGCTCGGATACACCGAGAGATCATGAACCCCGACTACACCAAGCTGACCATGG 850

DB 154 CATACAGCTCGGATACACCGAGAGATCATGAACCCCGACTACACCAAGCTGACCATGG 213

QY 851 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACCAAGCTCCCTGCCAGCATCAGTACCT 910

DB 214 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACCAAGCTCCCTGCCAGCATCAGTACCT 273

QY 911 TCGTGGAGGGCTACTCGAGCACTACGAATCAAGCTTCTCGGTGTACCAATGCGAGC 970

DB 274 TTGTGGAGGGCTACTCGAGCACTACGAATCAAGCTTCTCGGTGTACCAATGCGAGC 333

QY 971 GGCCTTGAATCAAGTGGAGGG 1030

DB 334 GGCCTTGAATCAAGTGGAGGG 393

QY 1031 ACCACACACCAACCAACATACACGAGCAGCATCAGCAGCATCCTTCTCCAGCT 1090

DB 394 ACCACACCAACCAACCAACATACACGAGCAGCATCAGCAGCATCCTTCTCCAGCT 453

QY 1091 CCAGCCCGGAGGACGAGGTGCTGCCAGCACCTTCCATGCTTCAAGCAGTCCCAACCGT 1150

DB 454 CCAGCCCGGAGGACGAGGTGCTGCCAGCACCTTCCATGCTTCAAGCAGTCCCAACCGT 513

QY 1151 CCACCCCCACCAACCGGGGCTTCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1210

DB 514 CCACCCCAACACGCGCGCTTCCCGCCGAGGGGGGGGTTATGGAGCAGGCACTGC 573

QY 1211 CCTCGGGCCCGGGCTGCATCGCACCCCGCGCTGCTGGAGCCCGCCCGATGAAGCGGTCC 1270

DB 574 CCTCGGGCCCGGGCTGCATCGCACCCCGCGCTGCTGGAGCCCGCCCGATGAAGCGGTCC 633

QY 1271 CCACGGTGGCGGGCGCGCTTCCCGCTTTCACATTTCAAGCCCTCGCGCGCATCCCC 1330

DB 634 CCACGGTGGCGGGCGCGCTTCCCGCTTTCACATTTCAAGCCCTCGCGCGCATCCCC 693

QY 1331 CCGCGCCAGCCCGGGCGGGCCACCACTCGGCTACGACCCAGCCGCGCTCGCGGC 1390

DB 694 CCGCGCCAGCCCGGGCGGGCCACCACTCGGCTACGACCCAGCCGCGCTCGCGGC 753

QY 1391 TCAGCCTGCGCTGGAGCGCGAGCCCGCGCGGAGCCAGCCCGCGCTTGGAGGCC 1450

DB 754 TCAGCCTGCGCTGGAGCGCGAGCCCGCGCGGAGCCAGCCCGCGCTTGGAGGCC 813

QY 1451 ACCGTACGGGCTGCGCTGGCCAAAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1510

DB 814 ACCGTACGGGCTGCGCTGGCCAAAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 873

QY 1511 TCAGCCTTCCCTTACCGCTTCCAGCTGCTGGGCGAGAGTCCAGAGCTGCGCTCGCGGC 1570

DB 874 TCAGCCTTCCCTTACCGCTTCCAGCTGCTGGGCGAGAGTCCAGAGCTGCGCTCGCGGC 933

QY 1571 CCACGAGAGCTCGCTGCTGGCGAGGGCGAGCTGCTGGCGGGGAGCAACCGCGCT 1630

DB 934 CCACGAGAGCTCGCTGCTGGCGAGGGCGAGCTGCTGGCGGGGAGCAACCGCGCT 993

QY 1631 GCCAGCACTAGGGCGTGGAACTTGGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGC 1690

DB 994 GCCAGCACTAGGGCGTGGAACTTGGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGC 1053

QY 1691 AGAAATGCAAAATATGTTGCTGCTGCAATAAACTGCCAGTAGAGAGAGAGCTC 1750

DB 1054 AGAAATGCAAAATATGTTGCTGCTGCAATAAACTGCCAGTAGAGAGAGAGCTC 1113

QY 1751 GAAACCGATGTCACTACTGTGATTTCAAGAGTGTCTCAGTGTGGAGTGTAAAGAG 1810

DB 1114 GAAACCGATGTCACTACTGTGATTTCAAGAGTGTCTCAGTGTGGAGTGTAAAGAG 1173

QY 1811 TTGTCCGTACAGATAGTCTGAAAGGGAGAGAGTGTCTGCTCTCCCAATCTGCATGATGATG 1870

DB 1174 TTGTCCGTACAGATAGTCTGAAAGGGAGAGAGTGTCTGCTCTCCCAATCTGCATGATGATG 1233

QY 1871 CATTAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCTCCCAATCTGCATGATGATG 1930

DB 1234 CATTAACAGGAACCTTCTCAGCCCTTCTCAGCCCTTCTCCCAATCTGCATGATGATG 1293

QY 1931 CCCTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTTGATTTATTCAGATCTGTC 1990

DB 1294 CTCTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTTGATTTATTCAGATCTGTC 1353

QY 1991 CCATGACAGGCTGCTGCGAGGACAGATGCTGAGCATGTGCAAACTTCTACACCTCC 2050

DB 1354 CCATGACAGGCTGCTGCGAGGACAGATGCTGAGCATGTGCAAACTTCTACACCTCC 1413

QY 2051 TGACAGCTTCCATTTGATGATTCAGAGAGCTGGGCGAGAAAGATTCGGGGATTTACTGATC 2110

DB 1414 TGACAGCTTCCATTTGATGATTCAGAGAGCTGGGCGAGAAAGATTCGGGGATTTACTGATC 1473

QY 2111 TCCCCAAGAAAGATCAGACATTACTTATGTAATCAGCCCTTTTGGAGCTGTTTGTCTCA 2170

DB 1474 TCCCCAAGAAAGATCAGACATTACTTATGTAATCAGCCCTTTTGGAGCTGTTTGTCTCA 1533

QY 2171 GACTTTCATCAGGTCAACACTGCTGAGATAGTGTGCTTCTGCAATGAGCTTGTCC 2230

DB 1534 GACTTTCATCAGGTCAACACTGCTGAGATAGTGTGCTTCTGCAATGAGCTTGTCC 1593

QY 2231 TGCACTCAGCTTTCAGTGCCTTCTGCTGAGTGTGGGAGTGGCTCGACTCTATTAAAGACTTTT 2290

```
Db 1594 TGATCGACTTCAGTCCCTTCGTGGATTGGGAGTGGCTCGACTCTATTAAAGACTTTT 1653
Qy 2291 CCTTAAATTTGACAGACCTGAACTTGTATATCCAAAGCCTTAGCTGCCTGTCAGCACTGA 2350
Db 1654 CCTTAAATTTGACAGACCTGAACTTGTATATCCAAAGCCTTAGCTGCCTGTCAGCACTGA 1713
Qy 2351 GCATCATCACAGAAAGACATGGTTTAAAGAACCAAGAGAGTCCGAAGAGCTATGCAACA 2410
Db 1714 GCATGATCACAGAAAGACATGGTTTAAAGAACCAAGAGAGTCCGAAGAGCTATGCAACA 1773
Qy 2411 AGATCACAGCAAGTTTAAAGAACCAAGAGAGTCCGAAGAGCTATGCAACA 2470
Db 1774 AGATCACAGCAAGTTTAAAGAACCAAGAGAGTCCGAAGAGCTATGCAACA 1833
Qy 2471 CCAAGTCTCTGGTCCCTTGTAGAACTGAGGAAGATCTGCAACCTCGGGCTCCAGCGCA 2530
Db 1834 CCAAGTCTCTGGTCCCTTGTAGAACTGAGGAAGATCTGCAACCTCGGGCTCCAGCGCA 1893
Qy 2531 TCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTCACAAGCTCTTCC 2590
Db 1894 TCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTCACAAGCTCTTCC 1953
Qy 2591 TGGACACCTACCTTTCTAATCAGAGCAGTGGAGCAGTGAGCTGCCTCCTCTCCTAGCA 2650
Db 1954 TGGACACCTACCTTTCTAATCAGAGCAGTGGAGCAGTGAGCTGCCTCCTCTCCTAGCA 2013
Qy 2651 CCTGCTGCTAGCAGCAAGGATAGGTTTGGAAACCTATCATTTCTGCTCCTTCCCTTA 2710
Db 2014 CCTGCTGCTAGCAGCAAGGATAGGTTTGGAAACCTATCATTTCTGCTCCTTCCCTTA 2073
Qy 2711 AGAGGAAAGCAGCTCCCTGTAGAAAGCAAGACTTCTTTTCTCGCTCTTTCCTT 2770
Db 2074 AGAGGAAAGCAGCTCCCTGTAGAAAGCAAGACTTCTTTTCTCGCTCTTTCCTT 2133
Qy 2771 ACAACCTAAAGCCAGAAACCTTGCAGAGTATTTGTTGGGGTTGTGTTTATATATTAGGC 2830
Db 2134 ACAACCTAAAGCCAGAAACCTTGCAGAGTATTTGTTGGGGTTGTGTTTATATTAGGC 2193
Qy 2831 ATTGGGGATGGGGTGGAGGGGTTATAGTTTCATGAGGGTTTCTAAGAAATTCCTAAC 2890
Db 2194 ATTGGGGATGGGGTGGAGGGGTTATAGTTTCATGAGGGTTTCTAAGAAATTCCTAAC 2253
Qy 2891 AAAGCAGCTTTGGCAAACTATCCACGACGMAAAAGGATATATACTGTTTAA 2950
Db 2254 AAAGCAGCTTTGGCAAACTATCCACGACGMAAAAGGATATATACTGTTTAA 2313
Qy 2951 AACTCTTTCTGGGAATCCAAATTAGTTGCTTTGTTTAAACAAAGAACAGCAAGG 3010
Db 2314 AACTCTTTCTGGGAATCCAAATTAGTTGCTTTGTTTAAACAAAGAACAGCAAGG 2373
Qy 3011 GTTGTTCGCCAGGATAGTGTCTTAAAGATTGGTCCCTTGAAAAATATGCTTCTGTA 3070
Db 2374 GTTGTTCGCCAGGATAGTGTCTTAAAGATTGGTCCCTTGAAAAATATGCTTCTGTA 2433
Qy 3071 TCAAGGTACGTATGTGTGCAACAGCGCAGAACTTCCCTTTAAATTCCTTCTTCTT 3130
Db 2434 TCAAGGTACGTATGTGTGCAACAGCGCAGAACTTCCCTTTAAATTCCTTCTTCTT 2493
Qy 3131 TATTTTAAACAATGGTGAAGATGAGGATTACCTTACAAATCAGACATGGCAAAACAATA 3190
Db 2494 TATTTTAAACAATGGTGAAGATGAGGATTACCTTACAAATCAGACATGGCAAAACAATA 2551
Qy 3191 ATGGCTGTTGCTTCCATAAACAAGTGCATTTTAAAGTGTGCTTCTTAAAGTCTTG 3250
Db 2552 ATGGCTGTTGCTTCCATAAACAAGTGCATTTTAAAGTGTGCTTCTTAAAGTCTTG 2608
Qy 3251 TTTATTAACTCTCCTTTATTTCTATATGGAATATAAAGAGGAGCTCATGTAGCAATG 3310
Db 2609 T---TATTAATCTCTTATCTATATGCGCGAAATAAAGAGGAGGAGCTCATGTAGCAATG 2665
Qy 3311 ACAGCTTAATATCCCTAGCAGGGCTGTGTTCACTTCCCTGTCGATCCCTTCTGAGGTA 3370
Db 2666 ACAGCTTAATATCCCTAGCAGGGCTGTGTTCACTTCCCTGTCGATCCCTTCTGAGGTA 2725
```

```
Qy 3371 TGGCCCATCCAAAGACTTTTAGGCCATTTCTGTATGGAACAGATCCCTGCCCTGACTGTCC 3430
Db 2726 TGGCCCATCCAAAGACTTTTAGGCCATTTCTGTATGGAACAGATCCCTGCCCTGACTGTCC 2785
Qy 3431 AGCTATCTCTGAAAGTGGATCAGATTATAACTGGATTACATCTAACTGTTTGGTTGCT 3490
Db 2786 AGCTATCTCTGAAAGTGGATCAGATTATAACTGGATTACATCTAACTGTTTGGTTGCT 2845
Qy 3491 TCTATCAACCCACCACAGAGTTCCCTTAAACTTTCCTTTCAGTTATAGTAAGTCTGATAT 3550
Db 2846 TCTATCAACCCACCACAGAGTTCCCTTAAACTTTCCTTTCAGTTATAGTAAGTCTGATAT 2904
Qy 3551 TCATTCAAGAGGCGCATTAAGTCAGTTGAGTATTTGATCCCTAGATAAGAACATGCAATC 3610
Db 2905 TCATTCAAGAGGCGCATTAAGTCAGTTGAGTATTTGATCCCTAGATAAGAACATGCAATC 2959
Qy 3611 AGCAGGAACCTGTCATACAGGCTAGCACAGGACCAATAAGGATTTTATAGATATAAT 3670
Db 2960 TCAGAGGACTGTCATACAGGCTAGCACAGGACCAATAAGGATTTTATAGATATAAT 3019
Qy 3671 TTAATTTTGTATTGTTTAAAGGACAAATTTTGGAGACAGCAAAA---TCTTTTAAA 3727
Db 3020 TTAATTTTGTATTGTTTAAAGGACAAATTTTGGAGACAGCAAAA---TCTTTTAAA 3079
Qy 3728 AAATAGTATGAATGTGAATCTAGAAAAGATTTAAAAATAGTATGAGTCTGAGTACTAG 3787
Db 3080 AAATAGTATGAATGTGAATCTAGAAAAGATTTAAAAATAGTATGAGTCTGAGTACTAG 3139
Qy 3788 GAAGGAT 3794
Db 3140 GAAGGAT 3146

RESULT 12
ADRI4062
ID ADRI4062 standard; DNA; 4977 BP.
XX AC ADRI4062;
XX AC ADRI4062;
XX DT 21-OCT-2004 (first entry)
XX DE Human NF-kappaB pathway-associated gene SeqID63.
XX DE NF-kappaB pathway; antinflammatory; cytostatic; hepatotropic; virucide;
KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;
KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
KW immunosuppressive; vulnery; gene therapy; immune disorder;
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia; immunodeficiency;
KW X-linked anhidrotic ectodermal dysplasia; hepatitis B; hepatitis C; EBV; influenza;
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW autoimmune disorder; hyper immune activity;
KW aberrant acute phase response; hypercongenital condition; birth defect;
KW necrotic lesion; wound; organ transplant rejection;
KW aberrant signal transduction; proliferating disorder; cancer;
KW HIV propagation; gene; ds; human.
XX OS Homo sapiens.
XX OS WO2004065577-A2.
XX PN WO2004065577-A2.
XX PD 05-AUG-2004.
XX PF 13-JAN-2004; 2004WO-US000798.
XX PR 14-JAN-2003; 2003US-0440068P.
XX PR 12-MAY-2003; 2003US-0469757P.
```

PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX Nadler SG, Neubauer MG, Feder JN, Carman J;
 PI WPI, 2004-562168/54.
 XX P-FSDB; ADR14063.
 DR
 DR
 XX
 PT New isolated polynucleotides and polypeptides associated with NF-kappaB
 PT pathway, useful for diagnosing, treating, or preventing disorders or
 PT diseases associated with NF-kappaB pathway.
 XX
 XX
 PS Claim 1; SEQ ID NO 63; 237pp; English.
 XX
 CC This invention relates to the novel association of protein sequences (and
 CC the genes which encode them) to the NF-kappaB pathway. The invention may
 CC be useful for the production of compounds with an antiinflammatory,
 CC cycostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
 CC gastrointestinal-gen, antilasthmatic, antiarteriosclerotic,
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnary activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions
 CC or diseases associated with the NF-kappaB pathway. The condition is an
 CC immune disorder, an inflammatory disorder, an inflammatory disorder
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
 CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
 CC immune activity, disorders related to aberrant acute phase responses,
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
 CC organ transplant rejection, conditions related to organ transplant
 CC rejection, disorders related to aberrant signal transduction,
 CC proliferating disorders, cancers and HIV propagation in cells infected
 CC with other viruses. The present sequence is that of a human gene which is
 CC subject to the novel association with the NF-kappaB pathway of the
 CC invention. Note: This sequence does not appear in the specification but
 CC was obtained by the indexer from Genbank.
 XX
 SQ Sequence 4977 BP; 1394 A; 1147 C; 1042 G; 1394 T; 0 U; 0 Other;
 Query Match 75.8%; Score 2875; DB 13; Length 4977;
 Best Local Similarity 98.1%; Pred. No. 0;
 Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;

Db 454 CCAGCCCGAGGACGAGGTGCTGCCAGACCTCCATGTACTTCAAGCAGTCCCAACCGT 513
 Qy 1151 CCACCCCCACACCGCCCGGCTTCCCGCCCGAGCGGGGGCGGTATATGGACAGGACACTGC 1210
 Db 514 CCACCCCCACACCGCCCGGCTTCCCGCCCGAGCGGGGGCGGTATATGGACAGGACACTGC 573
 Qy 1211 CTTGCGCGCGCGGTGATCGACCCCGCGCGCTGCTTGGACCCCGCGGATGAAGGGGTCC 1270
 Db 574 CTTGCGCGCGCGGTGATCGACCCCGCGCGCTGCTTGGACCCCGCGGATGAAGGGGTCC 633
 Qy 1271 CCAGGTGGCGCGCGGTGCTTCCGCTCTTCCACTTCAAGCCCTCCGCGCGCATCCCC 1330
 Db 634 CCAGGTGGCGCGCGGTGCTTCCGCTCTTCCACTTCAAGCCCTCCGCGCGCATCCCC 693
 Qy 1331 CCGCGCCAGCCCGCGCGCGGACCACTCGGCTACGACCCAGCCGCGCTGCGCGC 1390
 Db 694 CCGCGCCAGCCCGCGCGGCGCACCACTCGGCTAGGACCCAGCGCGCTGCGCGC 753
 Qy 1391 TCAGCTGCGCTGGGAGCCGACCGCGCGGCGAGCGAGCGCGCGCTTGAAGCC 1450
 Db 754 TCAGCTGCGCTGGGAGCCGACCGCGCGGCGAGCGAGCGCGCGCTTGAAGGCC 813
 Qy 1451 ACCGTAGGGCTGCGCTGCGCAAGAGGGCGCGCGCTGCGCTTCCGCGCTCTCGGC 1510
 Db 814 ACCGTAGGGCTGCGCTGCGCAAGAGGGCGCGCGCTGCGCTTCCGCGCTCTCGGC 873
 Qy 1511 TCAGCGCTTCCCTTACCGCGCTCAGACCTGTGGCGGAGAGTCCAGCGCTGCGCGC 1570
 Db 874 TCAGCGCTTCCCTTACCGCGCTCAGACCTGTGGCGGAGAGTCCAGCGCTGCGCGC 933
 Qy 1571 CCAGCAGAGTCTGCTGTGGCGAGGCGAGTGTGCGGTGTGGGGGCAACCGCGCT 1630
 Db 934 CCAGCAGAGTCTGCTGTGGCGAGGCGACGCTGTGCGGTGTGGGGGCAACCGCGCT 993
 Qy 1631 GCCAGCAGTCTGCGGTGCGACCTGCGAGGCTGCAAGGGCTTTTCAAGAGAGACAGTGC 1690
 Db 994 GCCAGCAGTCTGCGGTGCGACCTGCGAGGCTGCAAGGGCTTTTCAAGAGAGACAGTGC 1053
 Qy 1691 AGAAAAATGCAAAATATGTTTGCCTGCGCAATAAAAACTGCCAGTAGTACAGAGAGCTC 1750
 Db 1054 AGAAAAATGCAAAATATGTTTGCCTGCGCAATAAAAACTGCCAGTAGTACAGAGAGCTC 1113
 Qy 1751 GAAACCGATGTCAGTCTGTCGATTTGAGAGTGTCTCAGTGTGGATGTGTAAGAGAG 1810
 Db 1114 GAAACCGATGTCAGTCTGTCGATTTGAGAGTGTCTCAGTGTGGATGTGTAAGAGAG 1173
 Qy 1811 TTGTCCGTACAGATAGTCTGAAAGGAGGAGAGTGTCTGCTTCCAAACCAAGAGGCC 1870
 Db 1174 TTGTCCGTACAGATAGTCTGAAAGGAGGAGAGTGTCTGCTTCCAAACCAAGAGGCC 1233
 Qy 1871 CATTACAAAGAGAACCTTCTCAGCCCTCTCCACCTTCTCTCCAATCTGCATGATGAATG 1930
 Db 1234 CATTACAAAGAGAACCTTCTCAGCCCTCTCCACCTTCTCTCCAATCTGCATGATGAATG 1293
 Qy 1931 CCCTTGTCCGAGCTTTAAGACCTCAACCCAGAGATCTTGATTTATTCAGATAGTCTC 1990
 Db 1294 CTCTTGTCCGAGCTTTAAGACCTCAACCCAGAGATCTTGATTTATTCAGATAGTCTC 1353
 Qy 1991 CCAGTACAGGCTGCTGCGAGGACAGATGCTGAGCATGTGCAACAAATCTTACAACTCC 2050
 Db 1354 CCAGTACAGGCTGCTGCGAGGACAGATGCTGAGCATGTGCAACAAATCTTACAACTCC 1413
 Qy 2051 TGACAGCCTCCATTTGATGATTCAGAGAGCTGGGCGAGAAAGATTCGGGATTTACTGATC 2110
 Db 1414 TGACAGCCTCCATTTGATGATTCAGAGAGCTGGGCGAGAAAGATTCGGGATTTACTGATC 1473
 Qy 2111 TCCCAAGAGATGACAGCATTTGATTTGATCAGCCCTTTTGGAGCTGTTTGTCTCA 2170
 Db 1474 TCCCAAGAGATGACAGCATTTGATTTGATTTGATTTGAGCTGTTTGTCTCA 1533
 Qy 2171 GACTTTCCATCAGGTCAAAACACTGCTGAAGATAAGTTTGTCTGCAATGGACTTTGTCC 2230
 Db 1534 GACTTTCCATCAGGTCAAAACACTGCTGAAGATAAGTTTGTCTGCAATGGACTTTGTCC 1593

QY 2231 TGCATCGATTTCAGTGCCTTCGTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTT 2290
DB 1594 TGCATCGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTT 1653
QY 2291 CCTTAAATTTGCAGAGCCTGAACTCTGATATCCAAAGCCTTAGCCTGCCTGTGACAGACTGA 2350
DB 1654 CCTTAAATTTGCAGAGCCTGAACTCTGATATCCAAAGCCTTAGCCTGCCTGTGACAGACTGA 1713
QY 2351 GCATGATCACAGAAAGACATGGTTTAAAGAA CCAAGAGAGTGCAGAGACTATGCAACA 2410
DB 1714 GCATGATCACAGAAAGACATGGTTTAAAGAA CCAAGAGAGTGCAGAGACTATGCAACA 1773
QY 2411 AGATCACAGACGATTTAAAGACCAACCAAGAGTAAAGGACAGGCTCTGAGGCCACCGAGT 2470
DB 1774 AGATCACAGACGATTTAAAGACCAACCAAGAGTAAAGGACAGGCTCTGAGGCCACCGAGT 1833
QY 2471 CCAAGGTCCTGGTGCCCTGGTGAAGATCAGGAAGATCTGCACCCCTGGGCTCCAGCGCA 2530
DB 1834 CCAAGGTCCTGGTGCCCTGGTGAAGATCAGGAAGATCTGCACCCCTGGGCTCCAGCGCA 1893
QY 2531 TCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCAATGACAAGCTCTTCC 2590
DB 1894 TCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCAATGACAAGCTCTTCC 1953
QY 2591 TGGACACCTTACCTTCTTAATCAGAGCAGTGGAGCAGTGCCTCTCTCTAGCA 2650
DB 1954 TGGACACCTTACCTTCTTAATCAGAGCAGTGGAGCAGTGCCTCTCTCTAGCA 2013
QY 2651 CTGCTCTCTACGCAAGAGGATAGTTTGGAAACCTTATCATTTCTGCTCTTCCCTTA 2710
DB 2014 CTGCTCTCTACGCAAGAGGATAGTTTGGAAACCTTATCATTTCTGCTCTTCCCTTA 2073
QY 2711 AGAGAAAGCAGCTCTGTGAGAAAGCAAGACTTTCTTTTTTCTGGCTCTTTTCCCTT 2770
DB 2074 AGAGAAAGCAGCTCTGTGAGAAAGCAAGACTTTCTTTTTTCTGGCTCTTTTCCCTT 2133
QY 2771 ACAACCTTAAGCCAGAAAACTTGCAGAGTATGTGTGGGGTGTGTTTTATATTAGGC 2830
DB 2134 ACAACCTTAAGCCAGAAAACTTGCAGAGTATGTGTGGGGTGTGTTTTATATTAGGC 2193
QY 2831 ATTGGGGATGGGTGGGAGGGGTATAGTTTCATGAGGGTTCCTAAGAAATTCCTAAC 2890
DB 2194 ATTGGGGATGGGTGGGAGGGGTATAGTTTCATGAGGGTTCCTAAGAAATTCCTAAC 2253
QY 2891 AAAGCAGCTTTTGGCAATGCTATCCAGCAGGAAAAAAGGATAATAACTGTGTTTAA 2950
DB 2254 AAAGCAGCTTTTGGCAATGCTATCCAGCAGGAAAAAAGGATAATAACTGTGTTTAA 2313
QY 2951 AACTCTTTCTGGGGAATCCAAATATAGTTGCTTTGATTTTAAACAAGAACAGCCAAAG 3010
DB 2314 AACTCTTTCTGGGGAATCCAAATATAGTTGCTTTGATTTTAAACAAGAACAGCCAAAG 2373
QY 3011 GTTGTTCGCCAGGTPAGGATGTCTTAAAGATGCTCCCTTGAATAATGCTTCCCTGA 3070
DB 2374 GTTGTTCGCCAGGTPAGGATGTCTTAAAGATGCTCCCTTGAATAATGCTTCCCTGA 2433
QY 3071 TCAAAGGTACGTATGTGGTGCAAAACAAGCAGAGAACTTCTTTTTAAATTTCTTCTCCTT 3130
DB 2434 TCAAAGGTACGTATGTGGTGCAAAACAAGCAGAGAACTTCTTTTTAAATTTCTTCTCCTT 2493
QY 3131 TATTTTAAACAATGGTGAAGATGAGGATTAACCTACAAATCAGACATGGGCAAAACAATA 3190
DB 2494 TATTTTAAACAATGGTGAAGATGAGGATTAACCTAC -AATCAGACATGGGCAAAACAAT- 2551
QY 3191 ATGGCTGTTTCTTCCATTAACAAGTCAATTTTTTAAAGTCTCTTACTAGTCTTG 3250
DB 2552 ATGGCTGTTTCTTCC - - -ATACACGCTCAATTTTTTAAAGTCTCTTACTAGTCTTG 2608
QY 3251 TTTTAAACTCTCCTTTATCTTATATGAAATAAAAGAGGAGCAGTCTATGTTAGCAATG 3310
DB 2609 T - - -TATACTCTTATCTATATGCGGAATAAAAGAGGAGCAGTCTATGTTAGCAATG 2665

QY 3311 ACAGGTTAATATCCCTAGCAGAGGCTGTGTTACCTTCCCTGTCCATCCCTTCTGAGTA 3370
DB 2666 ACAGGTTAATATCCCTAGCAGAGGCTGTGTTACCTTCCCTGTCCATCCCTTCTGAGTA 2725
QY 3371 TGGCCCATCCCAAGACTTTTAGGCCATTCTTGATGGAACAGATCCCTGCGCTGACTGTCC 3430
DB 2726 TGGCCCATCCCAAGACTTTTAGGCCATTCTTGATGGAACAGATCCCTGCGCTGACTGTCC 2785
QY 3431 AGCTATCCTGAAAAGTGGATCAGATTATAAACTGGATTACATGTAACCTGTTTGGTGTGT 3490
DB 2786 AGCTATCCTGAAAAGTGGATCAGATTATAAACTGGATTACATGTAACCTGTTTGGTGTGT 2845
QY 3491 TCTATCAACCCCAACAGAGTTCCTTAAACTTGCCTTACCTTATAGTAACTGACTGTATAT 3550
DB 2846 TCTATCAACCCCAACAGAGTTCCTTAAACTTGCCTTACCTTATAGTAACTGACTGTATAT - 2904
QY 3551 TCATTTCAGAAGCGCCATAAAGTCAGTTGAGTATTTGATCCCTAGATAAAGCAATGCAAAATC 3610
DB 2905 TCATTTCAGAAGCGCCATAAAGTCAGTTGAGTATTT - -GATCCTAGATAAAGCAATGCAA --- 2959
QY 3611 AGCAGGAACCTGGTCATACAGGTAAGCACAGGGAACAATAAGGATTTTATAGATATAAT 3670
DB 2960 TCAGAGGACTGGTCATACAGGTAAGCACAGGGAACAATAAGGATTTTATAGATATAAT 3019
QY 3671 TTAATTTTGTATTGTTTAAAGGACAAATTTTGGAGCAAGCAAA - - -TCTTTTAAA 3727
DB 3020 TTAATTTTGTATTGTTTAAAGGACAAATTTTGGAGCAAGCAAAATCTTCTTTTAAA 3079
QY 3728 AAATAGTATGAATGTGAATACCTAGAAAAAGATTTAAAAAATAGTATGAGTGTGACTACTAG 3787
DB 3080 AAATAGTATGAATGTGAATACCTAGAAAAAGATTTAAAGAAATAGTATGAGTGTGACTACTAG 3139
QY 3788 GAAGGAT 3794
DB 3140 GAAGGAT 3146
RESULT 13
ADI31602
ID ADI31602 standard; cDNA; 2714 BP.
XX
AC ADI31602;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human cDNA #928.
XX
KW Human; gene; ss; immunological response; immunopathological condition;
KW Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
KW irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW acute monocytic leukaemia; antiinflammatory; antiaethmatic; antiulcer;
KW osteopathic; antiarthritic; antirheumatic; cytostatic.
XX
OS Homo sapiens.
XX
FN US6607879-B1.
XX
PD 19-AUG-2003.
XX
PF 09-FEB-1998; 98US-00023655.
XX
PR 09-FEB-1998; 98US-00023655.
XX
PA (INCY-) INCYTE CORP.
XX
PI Cocks BG, Stuart SG, Seilhamer JJ;
XX
XX
DR WPI; 2003-895307/82.
XX
PT A composition comprising a plurality of cDNAs, useful for detecting
PT altered expression of genes in an immunological response or for
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
PT or osteoarthritis.

XX Claim 1; SEQ ID NO 928; 50pp; English.

PS The invention relates to a composition comprising a plurality of cDNAs

XX for detecting the altered expression of genes in an immunological

CC response. The invention also relates to a method of diagnosing or

CC monitoring the treatment of an immunopathological condition in a sample,

CC comprising obtaining nucleic acids from a sample, contacting the nucleic

CC acids of the sample with an array comprising the plurality of cDNAs under

CC conditions to form one or more hybridisation complexes, detecting the

CC hybridisation complexes and comparing the levels of the detected

CC hybridisation complexes with the level of hybridisation complexes

CC detected in a non-diseased sample, where an altered level of the detected

CC hybridisation complexes correlates with the presence of an

CC immunopathological condition. Also disclosed are an expression profile

CC comprising a microarray and a plurality of detectable complexes and a

CC method for identifying a plurality of polynucleotide probes. The cDNAs

CC are useful as hybridisable array elements in a microarray for monitoring

CC the expression of target polynucleotides. The microarray can be used in

CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,

CC ulcerative colitis, hyperosinophilia, irritable bowel syndrome,

CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in

CC identifying agents for the treatment of the diseases. The microarray may

CC also be used in drug discovery and development, toxicological and

CC carcinogenicity studies, forensics or pharmacogenomics. The composition

CC may also be used in purification of a subpopulation of mRNAs, cDNAs or

CC genomic fragments. This sequence represents a human cDNA of the

CC invention. Note: The sequence data for this patent did not form part of

CC the printed specification but was obtained in electronic format directly

CC from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 2714 BP; 596 A; 924 C; 680 G; 514 T; 0 U; 0 Other;

Query Match 70.9%; Score 2689.6; DB 11; Length 2714;

Best Local Similarity 99.9%; Fred. No. 0;

Matches 2692; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 132 TCACCTCGCACACAGACACAGCGCGCACACAGGCTCCGACACACACTTGGCTCTCCC 191

DB 19 TCACACAAACACACAGACACAGCGCGCACACAGGCTCCGACACACACTTGGCTCTCCC 78

QY 192 CGCGGCTCACACCCCTCTTGGCCCTGAGCCCTTGGCGGTGACGCGCGCGCGAGCTGGA 251

DB 79 CGCGGCTCACACCCCTCTTGGCCCTGAGCCCTTGGCGGTGACGCGCGCGCGAGCTGGA 138

QY 252 CGCCCTCCCGGGCTCAGTTTGCACGCTGAGCGTGGCGGAGTGGCGGTGGAGTGGGA 311

DB 139 CGCCCTCCCGGGCTCAGTTTGCACGCTGAGCGTGGCGGAGTGGCGGTGGAGTGGGA 198

QY 312 ACAGCGCGGATCTCCCTCCCTTGGTCAAGCCCAAGCCAGGACGCGCGCGGAACCTCTC 371

DB 199 ACAGCGCGGATCTCCCTCCCTTGGTCAAGCCCAAGCCAGGACGCGCGCGGAACCTCTC 258

QY 372 GGCTGTGCTTCCATGATGCTGGATCGAGATCCCGACAGCGGCTGACCGGCTCG 431

DB 259 GGCTGTGCTTCCATGATGCTGGATCGAGATCCCGACAGCGGCTGACCGGCTCG 318

QY 432 GGAGCGGCTGGGCTTGTACACGCGAGCCCTTCCGGGACAGAGCTGTGACCTCCCGCCAG 491

DB 319 GGAGCGGCTGGGCTTGTACACGCGAGCCCTTCCGGGACAGAGCTGTGACCTCCCGCCAG 378

QY 492 TGCAGATTTGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAACCGCCACAGCACTC 551

DB 379 TGCAGATTTGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAACCGCCACAGCACTC 438

QY 552 AAAGCCCACTGGGAAGAGGCGAGCCCGGACAGCCCGGCGCTGAGCTTGAACCTTAGC 611

DB 439 AAAGCCCACTGGGAAGAGGCGAGCCCGGACAGCCCGGCGCTGAGCTTGAACCTTAGC 498

QY 612 GGTGCGGCGAGCACTGCGGGGCTTGGCTCGCGGAGAGTGGCTTCTTACACTCTC 671

DB 499 GGTGCGGCGAGCACTGCGGGGCTTGGCTCGCGGAGAGTGGCTTCTTACACTCTC 558

QY 672 AGCCTCCGCTGGAGAGACCCCGAGCCCAACATTCAGCGCGCAAGATACCTCTCAGATAT 731

DB 559 AGCCTCCGCTGGAGAGACCCCGAGCCCAACATTCAGCGCGCAAGATACCTCTCAGATAT 618

QY 732 GCCCTGGCTCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGCGCAGAC 791

DB 619 GCCCTGGCTCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGCGCAGAC 678

QY 792 ATACAGCTCCGAATACACACAGGAGATCATGAACCCCGACTACACCAAGCTGACCATGGA 851

DB 679 ATACAGCTCCGAATACACACAGGAGATCATGAACCCCGACTACACCAAGCTGACCATGGA 738

QY 852 CTTTGGCAGCACTGAGATCAAGGCTACAGCCACCACTCCCTGCGCAGCATGACATCTT 911

DB 739 CTTTGGCAGCACTGAGATCAAGGCTACAGCCACCACTCCCTGCGCAGCATGACATCTT 798

QY 912 CGTGGAGGGTACTCGAGCACTACGAACCTTCAGGCTTCTGCGGTGACCAAAATGACGCG 971

DB 799 CGTGGAGGGTACTCGAGCACTACGAACCTTCAGGCTTCTGCGGTGACCAAAATGACGCG 858

QY 972 GCCCTTGATCAAAAGTGGAGAGGGCGGGCGCCAGCTACCATCACCATCACCAACCA 1031

DB 859 GCCCTTGATCAAAAGTGGAGAGGGCGGGCGCCAGCTACCATCACCATCACCAACCA 918

QY 1032 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1091

DB 919 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 978

QY 1092 CAGCGCGGAGGACAGGCTGCGCCAGCACCTTCCATGCTTCAAGAGTCCCGACCGTCC 1151

DB 979 CAGCGCGGAGGACAGGCTGCGCCAGCACCTTCCATGCTTCAAGAGTCCCGACCGTCC 1038

QY 1152 CACCCCAACCAACCGCGCTTCCCGCGAGCGGGGGCGGTTATGGGACGAGGCACTGCC 1211

DB 1039 CACCCCAACCAACCGCGCTTCCCGCGAGCGGGGGCGGTTATGGGACGAGGCACTGCC 1098

QY 1212 CTGCGGCGCGGCTGCAATGCGCACCCCGCGCGCTCTGACACCGCGGCTGAGAGCGGCTCC 1271

DB 1099 CTGCGGCGCGGCTGCAATGCGCACCCCGCGCGCTCTGACACCGCGGCTGAGAGCGGCTCC 1158

QY 1272 CACGCTGGCGCGGCGGCTTCCCGCTTCCACTCAAGCCCTCGCGCGCGCATCCCGCC 1331

DB 1159 CACGCTGGCGCGGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCGCATCCCGCC 1218

QY 1332 CGCGCCAGCT 1391

DB 1219 CGCGCCAGCT 1278

QY 1392 CAGCTTGGCGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1451

DB 1279 CAGCTTGGCGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1338

QY 1452 CCGCTGAGCGGCTCGCGTGGCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCTTCCCGCT 1511

DB 1339 CCGCTGAGCGGCTCGCGTGGCAAGAGGCGCGCGCGCGCGCGCGCGCGCTTCCCGCT 1398

QY 1512 CACGCTTCCCTTACCGGCTCCAGCCTGCTGGGCGAGAGTCCAGCCTGCGCTGCGCGCC 1571

DB 1399 CACGCTTCCCTTACCGGCTCCAGCCTGCTGGGCGAGAGTCCAGCCTGCGCTGCGCGCC 1458

QY 1572 CAGCAGAGCTCGCTGCTGCGGAGGCGAGCTGCTGCGGTGCGGGGAGCAAGCGCGCTG 1631

DB 1459 CAGCAGAGCTCGCTGCTGCGGAGGCGAGCTGCTGCGGTGCGGGGAGCAAGCGCGCTG 1518

QY 1632 CCAGCACTTACCGGCTGCGAACTCGAGGCGCTGCAAGGCGCTTTTCAAGAGAACAGTGA 1691

DB 1519 CCAGCACTTACCGGCTGCGAACTCGAGGCGCTGCAAGGCGCTTTTCAAGAGAACAGTGA 1578

QY 1692 GAAATAATGCAAAATATGTTTCCCTGGGCAATAAATACTGCCAGTGTAGACAGAGCGTGC 1751

DB 1579 GAAATAATGCAAAATATGTTTCCCTGGGCAATAAATACTGCCAGTGTAGACAGAGCGTGC 1638

QY 1752 AAA CGATGT CAGTACTGT CGATTTTCAAGAGTGTCT CAGTGTGGAGTGTAAAGAGT 1811


```
Db 1639 AAACCGATGTCAGTACTGTCGATTTTCAGAAAGTGTCTAGTGTGGAATGGTAAAGAAGT 1698
Qy 1812 TGTCCGTAACAGATAGTCTGAAAGGAGGAGAGGTGCTGCTGCTTCCAAACCAAGAGCCC 1871
Db 1699 TGTCCGTAACAGATAGTCTGAAAGGAGGAGAGGTGCTGCTGCTTCCAAACCAAGAGCCC 1758
Qy 1872 ATTACACAGAGAACTTCTGAGCCCTCTCCACCTTCTCCTCCATCTCATGATGAATGC 1931
Db 1759 ATTACACAGAGAACTTCTGAGCCCTCTCCACCTTCTCCTCCATCTCATGATGAATGC 1818
Qy 1932 CTTGTCCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTTATCCAGATCTGTGCC 1991
Db 1819 CTTGTCCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTTATCCAGATCTGTGCC 1878
Qy 1992 CACTGACAGGCTGCTGAGGACAGATGCTGAGCATGTGCAACAATTTCTACAACCTCCT 2051
Db 1879 CACTGACAGGCTGCTGAGGACAGATGCTGAGCATGTGCAACAATTTCTACAACCTCCT 1938
Qy 2052 GACAGCCTCCATTTGATGTATTCAGAGCTGGGAGAGAAAGATTCGGGATTTACTGATCT 2111
Db 1939 GACAGCCTCCATTTGATGTATTCAGAGCTGGGAGAGAAAGATTCGGGATTTACTGATCT 1998
Qy 2112 CCCAAAGAGATCAGACATTTACTTATTTGAATCAGCCTTTTGGAGCTGTTTGTCTCAG 2171
Db 1999 CCCAAAGAGATCAGACATTTACTTATTTGAATCAGCCTTTTGGAGCTGTTTGTCTCAG 2058
Qy 2172 ACTTTCCATCAGGTCACAACTGCTGAAAGATAAGTTTGTGTTCTGCAATGACCTTGTCT 2231
Db 2059 ACTTTCCATCAGGTCACAACTGCTGAAAGATAAGTTTGTGTTCTGCAATGACCTTGTCT 2118
Qy 2232 GCATCGACTTCAGTGCCCTTCTGGAATTTGGGAGTGCGACTCTATTAAGACTTTTC 2291
Db 2119 GCATCGACTTCAGTGCCCTTCTGGAATTTGGGAGTGCGACTCTATTAAGACTTTTC 2178
Qy 2292 CTTAAATTTGAGAGCCTGACCTTGATATCCAGCCTTTAGCCTGCTGTCAGCACTGAG 2351
Db 2179 CTTAAATTTGAGAGCCTGACCTTTGATATCCAGCCTTTAGCCTGCTGTCAGCACTGAG 2238
Qy 2352 CATGATCACAGAAAGACATGGGTTTAAAGAACCAAGAGAGTTCGAAGAGCTATGCAACA 2411
Db 2239 CATGATCACAGAAAGACATGGGTTTAAAGAACCAAGAGAGTTCGAAGAGCTATGCAACA 2298
Qy 2412 GATCAACAGAGCTTTAAAGAACCAAGAGTAAGGGAAGGCTCTGGAGGCCACCGAGTC 2471
Db 2299 GATCAACAGAGCTTTAAAGAACCAAGAGTAAGGGAAGGCTCTGGAGGCCACCGAGTC 2358
Qy 2472 CAAGGTCCTGGGTGCTGCTGAGTACTGAGAGATCTGCACCTTGGGCTTCCAGCGCAT 2531
Db 2359 CAAGGTCCTGGGTGCTGCTGAGTACTGAGAGATCTGCACCTTGGGCTTCCAGCGCAT 2418
Qy 2532 CTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATTCATTGACAAAGCTTCTCT 2591
Db 2419 CTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATTCATTGACAAAGCTTCTCT 2478
Qy 2592 GGACACCTTACCTTTTCTAATCAGGAGCAGTGGAGCAGTGAGTGCTCTCTCTTAGCAC 2651
Db 2479 GGACACCTTACCTTTTCTAATCAGGAGCAGTGGAGCAGTGAGTGCTCTCTCTTAGCAC 2538
Qy 2652 CTGCTTGCTAGCAGCAAGGATAGGTTTGGAAACCTATCATTTCTCTGCTTCTTCTTAA 2711
Db 2539 CTGCTTGCTAGCAGCAAGGATAGGTTTGGAAACCTATCATTTCTCTGCTTCTTCTTAA 2598
Qy 2712 GAGGAAAGAGAGCTCTCTAGAAAGCAAGACTTCTTTTCTTTTCTGCTCTTTTCTTAA 2771
Db 2599 GAGGAAAGAGAGCTCTCTAGAAAGCAAGACTTCTTTTCTTTTCTGCTCTTTTCTTAA 2658
Qy 2772 CAACCTAAGCCAGAAAATCTTGACAGATATTTGTTGGGGTGTGTTTATATTTA 2827
Db 2659 CAACCTAAGCCAGAAAATCTTGACAGATATTTGTTGGGGTGTGTTTATATTTA 2714
```

RESULT 14

AAT16151
ID AAT16151 standard; cDNA to mRNA; 4400 BP.
XX
AC AAT16151;
DT 19-JUL-1996 (first entry)
XX
DE Apoptotic cerebral neuron nuclear receptor gene.
XX
KW Rat nuclear receptor gene; embryo; cerebral neuron; induction; apoptosis;
KW amplification; primer; PCR; Antisense; ribozyme; nerve; disease; ss.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT CDS 700..2586
FT /*tag= a
FT /product= "rat nuclear receptor"

JP08023980-A.

30-JAN-1996.

15-JUL-1994; 94JP-00164434.

15-JUL-1994; 94JP-00164434.

(TERU) TERUMO CORP.

WPI; 1996-133421/14.

P-PSDB; AAR92057.

Nuclear receptor expressed in apoptosis of cerebral neuron - used to
design anti-sense oligo:nucleotide(s) and ribozyme(s) for treatment of
cerebral nerve disease etc.

Claim 1; Page 7-10; 11pp; Japanese.

This is the nucleotide sequence of the novel rat nuclear receptor gene
designated NUR88. The gene was isolated from 16 day old rat embryo
cerebral neurons induced into apoptosis by amplifying the sequence from
cDNA using primers AAT28371-4. Antisense nucleic acids and ribozymes
targeted to the gene can be used as agents for treating cerebral nerve
diseases

Sequence 4400 BP; 1055 A; 1239 C; 1034 G; 1072 T; 0 U; 0 Other;

Query Match 65.9%; Score 2500; DB 2; Length 4400;

Best Local Similarity 83.4%; Pred. No. 0;

Matches 3131; Conservative 0; Mismatches 535; Indels 90; Gaps 22;

Qy 46 CGGAGTCTCTGCTTCCGCGCCGCCACCCCTCCAGCTCCTGCTCCTCCGCTCCCAT 105

Db 1 CCGAGTCTCTGCTTCCGCGCCGCCACCCCTCCAGGCTGCTCCTCCTCCGCTCCCAT 60

Qy 106 ACACAGAGCGCTCACACCCGCTCCTCTCATCTGCACACACAGACGCGGCTCACAG 165

Db 61 ACACAGACAGCTCACACCCGCTCCTTCACTTGACACACAGACGCGGCTCACAC 120

Qy 166 GCTCGGACACACACT--TCGCTCTCCGCGCGGCTCACACCCCTTTCGCTCCGCTT 223

Db 121 GCTCGGACACACACTTCCACTCTCTCCGCGCGGCTCACACCCCTCTCTCGGCGCC 180

Qy 224 GCGGCT-----GCAGCGCGCGCGCGAGCTGAGCGCCCTCCGCGCTCCTTTCGCAAG 278

Db 181 GCCGCTGTCGCGCGCGCGCGCGAGCGCCCTCAGGGCTCCTCCTTTCGCAAG 240

Qy 279 CTGACGCTGCGCGGAGTGGCGTGGAGGTGGGAAACAGCGGCGGCTCCTCCGCTTGGTC 338

Db 241 CTGACAGAGCGGCGAGTGGCGTGGAGGTGGGAAACGTTGGGACATCTTAGCCCTGGTC 300

Qy 339 ACAGCCCAAGCCAGAGCGCCCGGGAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATC 398

Db 301 GCAGCCGAGACTGGAGC-CTGCGGAACCTCTCGCGCGCTCTCCCATGATGTTGGATC 359
Qy GCAGATCCCCACACAGCC---GTTACCGGCTCGGGAGCGCTGGGCTTGTATACACGC 455
Db 360 GCAGATCCCGAGCAGCGCTGTCTACCGGCTCTGGAGCGCTGGGTTTGTGACCGC 419
Qy 456 AGCCCTTCGGGACAGCAGCTGTGATCTCCCCCAGTGCAGATTTCCGGGACAGCTCTCTA 515
Db 420 AGCCCTTCGGGACAGCAGCTGTGACTCTCCCCAATCCAGATTTCCGGGCTCGCTCTCTA 479
Qy 516 GAAACTGCTCTAAGACGGAAACCGCACAGCACTCAAAGCCCACTCGGGAAGGGGAG 575
Db 480 GAAACTGCTCTAAGACGGAAACCTCCACAGAACCCCAAAGCCCACTCGGGAAGGGGAG 539
Qy 576 CCGGCAAGCCCGGGCCCTGAGCTTGGACCTTATAGCGTGC CGGCGACGA-----CTGC 629
Db 540 CCGCAAGCCCGGGCGCTGAGCTTGGACCTTGAACCTTCAAGAGCGGCGACGACGGCGGC 599
Qy 630 CGGCGCTTTCGCTTCGCGGAGCTTCGCTCTCTCTACCTCTCAGCTCCGCTGGAGAGAC 689
Db 600 GGCTGCTTCGCTTATCCGAGCTCCCGGCTCTCTACACTCTCAGCTCCGCTGGAGAGAC 659
Qy 690 CCCAGCCCCACCAATTCAGCGCGCAAGATACCTTCCAGATATGCGCTGCTCCAGGCCA 749
Db 660 CCCAGCCCCACCAATTCAGCGCGCAAGATACCTTCCAGATATGCGCTGCTGCAAGGCCA 719
Qy 750 ATATAGCCCTTCCCTTCAGGTTCCAGTTATGCGGCGCAGACATACAGCTCGGAATACAC 809
Db 720 ATATAGCCCTTTCGCTTCGCGGCTTCATTTATGCGCAGCACTTATGGCTCGGAATACAC 779
Qy 810 CACGGAGATCATGAACCCCGACTACACCAAGCTGACCATGGAACCTTGGCAGCACTGAGAT 869
Db 780 CACAGAAATCATGAACCCCGACTATGCAAGCTGACCATGGAACCTTGGTACACCGGGAT 839
Qy 870 CACGGCTACGCCACAGCTCCCTGCCAGCATCAGTACCTTCTGGAGGCTACTCGAG 929
Db 840 CATGCCACGCCACAGCTCCCTGCCAGCTTCAGTACCTTTCATGGAGGGCTTACCCACG 899
Qy 930 CAATACGAACTCAAGCTTCTCGGTGTACCAATGCGAG-----CGGCCCTT 977
Db 900 CAGCTCGAACTCAAGCCCTCTGCTGTACCAATGCGCGCTTCTGGGCGCTCGGCCCTT 959
Qy 978 GATCAAAAGTGGAGGGGCGGCGCCAGCTACCATCAACCATCAACCAACCAACCA 1037
Db 960 GATCAAGATGGAAGGGTCCGAGCATGGCTTACCAACCAACCAACCAACCAACCA 1019
Qy 1038 CCACACCAACCATCACAGCAGCAGCATCAGCAGCATCCATTCCTCAGGCTCCAGGCC 1097
Db 1020 TCATCACCAACCAACCAAGCAG-----CAGCAGCGCTCCATTCCTCCTCCTCGGCC 1073
Qy 1098 GGAGACGAGGTGTGCCACGACCTCCATGTACTTCAAGCAGTCCCAACCGCTCCACCCC 1157
Db 1074 CGAGGACGAGGTATGCGCCAGCACTCATGTATCTTCAAGCAGTCTCGCGGCTTACGCC 1133
Qy 1158 CACACGCGGCTTTCGCCCGCAGGGGCGGCGTATGGACGAGGCACTGCCCTCGGC 1217
Db 1134 GACCCTCCAGGCTTCCCCCGCAGGGGCGGCGCTGTGGACGACGAGCTGCCCTCTGC 1193
Qy 1218 GCGCGCTGATCGACCCCGCGCTGCTGTGACCCCGCGATGAAGGGGTTCCCAAGGT 1277
Db 1194 GCTTGGCTGATCGCTCCGGGACCGCTGCTGGAACCGCAGATGAAGGCGAGTGCCTCCAT 1253
Qy 1278 GCGCGCGCGCTTTCGCGCTTTCCTCACTTCAAGCCCTTCGCGCGCATCCCCCGGCC 1337
Db 1254 GCGCGCTGCTCGCGCTTCCCGATCTTCTTCAAGCCCTTACCGCAACCTTCCCGGCC 1313
Qy 1338 CAGCCCGCGCGCGCCACCACTCTGGCTACGACCCGACCGCGCTGCGCGCTCAGCCT 1397
Db 1314 CAGCCACGCGCGCGCCACCACTCTGGGCTATGACCCCAACCGCGCGCTGCGCTCAGTCT 1373
Qy 1398 GCGCTGGAGCGCAGCGCGCGCGGAGCAGCGCGCGCTTGAAGCAGCCCGTA 1457
Db 1374 ACCCTGGAGCGCGCGCGCGCGGAGCAGCAAGCTGCTGCGCTCGAGGGCCATCCGTA 1433

Qy 1458 CGGCTGCGCTGCGCCAAAGAGGGCGGCCGCTGCGCTTCCGCGCTCTCGGCTCACGCC 1517
Db 1434 CGGCTGCGCTGCGCCAAAGAGGACGGCCACGTTGACCTTCCCTCGCTGGGCTCACAGC 1493
Qy 1518 CTCCCTTACCGCTGCGCAGCTGTGGGCGAGGTCCAGAGCTGCGCTGCGCGCCAGAG 1577
Db 1494 GTCCCTTACCGCTGCGCAGCTGTGGGAGAGAGCCCGCTACCATCGCCACCAATAG 1553
Qy 1578 GAGTCTGCTGCTGCGCGAGGGCAGCTGCGCTGCGGGGACAAACGCGCTGCCAGCA 1637
Db 1554 GAGTCTCATCATTCGCGCGAGGGCAGCTGTGCTGTGTGCGGGGCAATGTGCTGCCAGCA 1613
Qy 1638 CTACGGCTGCGAACTCTCGAGGGCTCAAGGGCTTTTTCAAGAGAAACAGTGCAGAAAAA 1697
Db 1614 CTACGGAGTCCGCACTTCCGAGGGCTCGAAGGGCTTTCTTCAAGAGAACGGTGCAGAAAAA 1673
Qy 1698 TGCAAAATATGTTTTCGCTGCGCAATAAATACTGCCAGTAGACAAGAGACGTCGAAACCG 1757
Db 1674 CGCAAAATATGTTTGTGCTTGGCAATAAATACTGCCAGTAGACAAGAGACGTCGAAATCG 1733
Qy 1758 ATGTCAGTACTGTCGATTTTCAAGAGTGTCTCAGTGTGGATGTAAAGAGTGTGTCG 1817
Db 1734 ATGTCAGTACTGTCAGGTTTTCAGAAAGTGTCTCAGTGTGCGGATGTGAAGAGTGTGCG 1793
Qy 1818 TACAGATAGTCTGAAAGGGGAGAGGTCTGCTGCTTCCAAACCAAGAGGCCCATTACA 1877
Db 1794 TACAGATAGTCTGAAAGGGGAGAGGTCTGCTGCTTCCAAACCAAGAGGCCCATTACA 1853
Qy 1878 ACAGGAACCTTTCAGGCCCTCTCCACCTTCTCCTCAATCTGCATGATGAATGCCCTGT 1937
Db 1854 ACAGGAGCCCTCGCAGGCCCTCCACCATCTCCTCCGATCTGTATGATGAACGCGCTTGT 1913
Qy 1938 CCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTTTCCAGATACTGCTCCACTGA 1997
Db 1914 CCGAGCTTTAACAGCCGACGCGCCAGAGACTTGTATTTCTCCAGATACTGCTCCACCGA 1973
Qy 1998 CCAGGCTGCTCGAGGCCACAGATGCTGAGCATGTGCAACAAATTTTACAACCTCTTGACAGC 2057
Db 1974 CCAGGCCACTCGCGGCACAGACGCTGAGCAGCTGACAGAGTTCTTACAACCTCTTGACGCGC 2033
Qy 2058 CTCCATTTGATGATTCAGAGAGCTGGGCAAGAAAGATTCCGGGATTTACTGATCTCCCAA 2117
Db 2034 CTCCATCGAGTGTCCAGAAAGCTGGGCAAGAAAGATCCCGGATTTCACTGATCTCCCAA 2093
Qy 2118 AGAAGATCAGACATTACTTATTTGAATCAGCCCTTTTGGAGCTGTTGTCTCAGACTTTC 2177
Db 2094 AGAAGATCAGACGTTACTTATAGAAATCAGCCCTTTTGGAGCTGTTGTCTTAGACTTTC 2153
Qy 2178 CATCAGGTCAAACACTGCTGAAAGATAAGTTTGTGTTTCTGCAATGGACTTGTCTGCAFCG 2237
Db 2154 TATCAGGTCAAACACTGCTGAAAGATAAGTTTGTGTTCTGCAATGGACTTGTCTGCAFCG 2213
Qy 2238 ACTTCAGTGCCTTCGTGGAATTTGGGAGTGGCTCGACTCTATTTAAAGACTTTTCTTAAA 2297
Db 2214 ACTTCAGTGCCTTCGCGAATTTGGGAGTGGCTCGACTTCCATTTAAAGACTTTTCTTAAA 2273
Qy 2298 TTTTCAGAGCTGAACCTTGAATCCAGGCTTAGCCTGCTGTCAGCAGCTGAGCATGAT 2357
Db 2274 TTTTCAGAGCTGAACCTTGAATCCAGGCTTAGCCTGCTGTCAGCAGCTGAGTATGAT 2333
Qy 2358 CACAGAAAGACATGGGTTAAAGAAACCAAGAGAGTGCAGAGAGCTATGCAACAAGATCAC 2417
Db 2334 CACAGAGGACATGGGTTAAAGAAACCAAGAGAGTGCAGAGAGCTATGCAACAAGATCAC 2393
Qy 2418 AAGCAGTTTAAAGACCAACAGAGTAAGGAGCAGGCTCTGGAGCCCAAGGTCCAAGGT 2477
Db 2394 AAGCAGCTTAAAGAGCACCAACAGAGTAAGGAGCAGGCTCTGGAGCCCTCAGAGCCCAAGGT 2453
Qy 2478 CTTGGGCTGCTGCTGAGAAAGATCTGACACCTTGGGCTTCCAGCGCATCTTCTTA 2537
Db 2454 CTTTCGCGCATGCTGTTGGAATCTGAGGAAGATCTGCACCCAGGCGCTTCCAGCGTATCTTCTTA 2513

Db 2154 TATCAGGTCAAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTTGTCCTGCACCG 2213
Qy 2238 ACTTCAGTGCCTTCGTGATTTGGGGAGTGGCTCGACTCTATTAAGAAGCTTTTCTTTAA 2297
Db 2214 ACTTCAGTGCCTTCGTGATTTGGGGAGTGGCTCGACTCTATTAAGAAGCTTTTCTTTAA 2273
Qy 2298 TTTGCAGAGCCTGAACCTTTGATATCAAGCCTTAGCCTGCTGTCAGACATGAGCATCAT 2357
Db 2274 TTTGCAGAGCCTGAACCTTTGATATCAAGCCTTAGCCTGCTGTCAGACATGAGCATCAT 2333
Qy 2358 CACGAAAGACATGGGTTTAAAGAACCAAGAGAGTTCGAAGAGCTTATGCAACAGATCAC 2417
Db 2334 CACAGAGCATGCTTTAAAGAACCAAGAGAGTTCGAAGAGCTTATGCAACAGATCAC 2393
Qy 2418 AAGCAGTTTAAAGACCAAGAGTAAGGACAGGCTCTGAGAGCCCAACGAGTCCAAGGT 2477
Db 2394 AAGCAGTTTAAAGACCAAGAGTAAGGACAGGCTCTGAGAGCCCAACGAGTCCAAGGT 2453
Qy 2478 CTTGGTGCCTGTGAGAACTGAGGAAGATCTGCACCTGGGCTCCAGCGCATCTTCTA 2537
Db 2454 CTTGGTGCCTGTGAGAACTGAGGAAGATCTGCACCTGGGCTCCAGCGCATCTTCTA 2513
Qy 2538 CTTGAAGCTGGAAGACTTGTGTCTCCACCTTCCATCATGACAAAGCTTCTCTGGACAC 2597
Db 2514 CTTGAAGCTGGAAGACTTGTGTCTCCACCTTCTCTCATGCAAGCTTCTCTGTATAC 2573
Qy 2598 CTTACCTTTTAATCAGAGCAG - TGGAGAGTGTGAGTGCCTCTCTAGCAGCTGCT 2656
Db 2574 CTTGCTTTCTGAGCAGGGAAGCTGAGCAGAGAGTACTTGTCTCTGCTGCACTGGTC 2633
Qy 2657 TGCTACGAGCAAGAGGATAGTTTGGAACTATCATCTTCTGCTCTCTTAAGAGGA 2716
Db 2634 ATTAAGTGAAGAAAGATGGTTTGAACACT - GCCCTCTCTATCTCTCTCCAGGGGAA 2692
Qy 2717 AAAGCAGCTCTGTAGAAAGCAAGACTTCTTTTCTGCTGCTCTTCTCTACAC 2776
Db 2693 AAAGCAGCTCTGTAGAAAGCAAGACTTCTTTTCTGCTGCTCTTCTCTACAC 2751
Qy 2777 TAAAGCCAGAAACTTGCAGAGTATGTGTGGGTGTGTGTATATTTAGGCAATTGGG 2836
Db 2752 TAAAGCCAGAAACTTGCAGAGTATGTGTGGGTGTGTGTATATTTAGGCTTTGGT 2811
Qy 2837 GGATGGGTGGAGGGGT - TATAGTTCTATGAGGTTTCTTAAGAAATGCTTAACAA 2894
Db 2812 GGATGGGTGGAGGGGTAAATAGTTCTATGAGGCTTTCTTAAGAAATGCTTAACAA 2871
Qy 2895 CACTTTTGGCAATGCTATCCAGCAG - GAAAAAAGGATATATAAAGTGT 2948
Db 2872 CACTTTTGGATGCTATCCAGCAGTGGGGTGGGGAAGGATATATAAAGTGT 2931
Qy 2949 AAAA - CTCTTTCTGGGGAATCCAAATATAGTGTCTTTGTATTTAAAAAAGCAAGCC 3006
Db 2932 AAAA - CTCTTTCTGGGGAATATGACTATGTTGCTTTGTATTTAAAAAAGCAAGCC 2991
Qy 3007 AAGGGTGT - TTTGGCAGGTAGGATGTCTTAAGATTTGCTCTGAAATATGCTTC 3065
Db 2992 AAGGGCTGT - TTTACAGGTAGGCTGTCTTAAGACTGATCCCTTTAGTATGCTTC 3051
Qy 3066 CTGTATCAAGGTAGTGTGTGTCAAACAGGCAAGAA - CTTTCTTT 3113
Db 3052 CCGGATC - GAGGCATTAAGTGTGCAATGAGCGGGGAAATCTTCTTCTTCAIT 3110
Qy 3114 TAAATTCCTTCTCTTTTAAATTAAGTGTGAAAGATGAGGATTTACCTACAAATCA 3173
Db 3111 CTTTCTTCTTCTTAAATTAAGTGTGCAATGAGCGGGGAAATCTTCTTCTTCAIT 3170
Qy 3174 GATAGGCAAAACAAATATAGTGTGCTTCTTCCATTAACAGTGCATTTTAAAGTGC 3233
Db 3171 GACTTAGCAAAATGATAATGAGTATTCGCTTCCATACATCAAGTGCATTTTAAAGTGC 3230
Qy 3234 TGTCTTACTAGTCTTGTATTAATTAAGTCTCTTCTTCTTATATGAAATAAAGGAGGC 3293
Db 3231 TGTCTTACTAGTCTTGTATTAATTAAGTCTCTCTCTCTTCTTATATGAAATAAAGGAGGC 3290

Qy 3294 AGTCATGTTAGCAAAATGACACGTTAAATATCCTAGCAGAGGCTGTGTTCACTTCCCTGT 3353
Db 3291 AGTCATGTTATCAAAACGGCGTCTCATTTTCTAGCTCACCTTGGTCCACCTGCCCTGT 3350
Qy 3354 CGATCCCTTCTGAGGTATGGCCCATCCAAAGACTTTTAGGCCATTTCTGATGGAACAGAT 3413
Db 3351 AGAACCTTTCGAGGTATGGCCCTTCTAAGACTTTTCAGGCCACTCTTGTGATGATTCGAC 3410
Qy 3414 CCC - TGCCTGACTGTCCAGCTATCCTGAAAGTGGATCAGATTTATAAAGT 3463
Db 3411 ACCCTCCCTCAACCCATGACTATCCAGATGTCTGATGGGATCAGGTTATAAAGT 3470
Qy 3464 GATTACATGTAACCTTTTGGTGTGT - TCTATCAACCCACAGAGTTCCCTAAAGT 3522
Db 3471 GATTGCAATGACTGTGTGCTGTGTGTTTGTCAACCTGGACAGAGTTCTCTAAACCTT 3530
Qy 3523 CTTCAAGTTATAGTAACCTGACTGGTATATTCAATTCAGAAAGC - GCCATAAGTCA 3575
Db 3531 CTTTAGTTGTAGCAAGTTCTGATTTCTCCATTCAGAAAGCCCAAGGAGCATTTGGGTGACT 3590
Qy 3576 TGAGTATTTGATCCTCTAGATTAAGAACATGCAAAATCAGCAGGAAGT - GGTCTATACAGGTA 3634
Db 3591 CGATCAAGGGTTAACCCTTAGGAGAACATGCAAAATAGTAGGAAGTGGTCTCAGACAGGTA 3650
Qy 3635 AGCACCGAGGCAATAAAGGATTTTATAGATAT - AATTTAATTTTGTATTGG 3687
Db 3651 AGCACCGAGATGATAAGGATTTATATAAATATATAAATAATTAATTTTGTATTGG 3710
Qy 3688 TTAAGGAGACAAATTTGGAGAGCAAGCAAAATCTTTT 3723
Db 3711 TTA - TAGACAAATTTGGAAAGCAAGAGATCATCT 3744

Search completed: March 21, 2005, 17:28:15
Job time : 1919 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: March 21, 2005, 11:57:11 ; Search time 599 Seconds

(without alignments)
10363.996 Million cell updates/sec

Title: US-10-608-863-1

Perfect score: 3794

Sequence: 1 ataataacgtcgagagaga.....gtgtgactactagaaggat 3794

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgm2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgm2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgm2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgm2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgm2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgm2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3779.4	99.6	3797	4	US-09-949-016-3789
2	3722	98.1	3802	4	US-09-949-016-905
3	2689.6	70.9	2714	4	US-09-023-655-928
4	1420.4	37.4	47184	4	US-09-949-016-12647
5	1420.4	37.4	47184	4	US-09-949-016-15531
6	497	13.1	3427	4	US-09-023-655-1400
7	497	13.1	3427	4	US-09-949-016-828
8	497	13.1	3427	4	US-09-949-016-4467
9	486.4	12.8	1797	3	US-09-277-078-1
10	363.8	9.6	2486	4	US-09-949-016-4392
11	340.6	9.0	2637	4	US-09-566-921-58
12	315.6	8.3	601	4	US-09-949-016-35444
13	315.6	8.3	601	4	US-09-949-016-135237
14	236	6.2	601	4	US-09-949-016-135236
15	174.6	4.6	601	4	US-09-949-016-35441
16	174.6	4.6	601	4	US-09-949-016-135227
17	135.6	3.6	601	4	US-09-949-016-35440
18	135.6	3.6	601	4	US-09-949-016-135226
19	109.6	2.9	1757	4	US-08-216-592A-3
20	109.6	2.9	2130	1	US-07-952-800-1
21	109.6	2.9	2538	4	US-09-949-016-5701
22	109.4	2.9	1245	2	US-09-132-619-3
23	109.4	2.9	1245	3	US-09-282-803B-3
24	109.4	2.9	1245	3	US-09-510-654-3
25	109.4	2.9	2330	2	US-09-132-619-7
26	109.4	2.9	2330	3	US-09-282-803B-7
27	109.4	2.9	2330	3	US-09-510-654-7

28	109.4	2.9	2482	4	US-09-949-016-548	Sequence 548, App
29	109.4	2.9	2486	4	US-09-949-016-1196	Sequence 1196, Ap
30	109.4	2.9	3115	2	US-09-132-619-1	Sequence 1, Appl
31	109.4	2.9	3115	3	US-09-282-803B-1	Sequence 1, Appl
32	109.4	2.9	3115	3	US-09-510-654-1	Sequence 5, Appl
33	109.4	2.9	3251	2	US-09-132-619-5	Sequence 5, Appl
34	109.4	2.9	3251	3	US-09-282-803B-5	Sequence 5, Appl
35	109.4	2.9	3251	3	US-09-510-654-5	Sequence 5, Appl
36	109.2	2.9	1330	4	US-10-329-668-17	Sequence 17, Appl
37	108.4	2.9	12232	4	US-09-949-016-12570	Sequence 12570, A
38	108.4	2.9	12240	4	US-09-949-016-16209	Sequence 16209, A
39	104.2	2.7	1564	4	US-09-949-016-1621	Sequence 1621, Ap
40	104.2	2.7	1827	4	US-10-329-668-15	Sequence 15, Appl
41	102.8	2.7	2204	1	US-07-952-800-3	Sequence 3, Appl
42	102.8	2.7	2219	1	US-08-336-408B-7	Sequence 7, Appl
43	102.8	2.7	2219	5	PCT-US91-00399-7	Sequence 7, Appl
44	102.8	2.7	2285	4	US-08-216-592A-1	Sequence 1, Appl
45	102.6	2.7	1594	4	US-09-016-434-1059	Sequence 1059, Ap

ALIGNMENTS

RESULT 1

US-09-949-016-3789
; Sequence 3789, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C001307
; CURRENT APPLICATION NUMBER: US/09/949.016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3789
; LENGTH: 3797
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3789

Query Match 99.6%; Score 3779.4; DB 4; Length 3797;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3793; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY	1	ATAAATGACGTGCGAGAGACGAGCGGAGCGGAGCGGAGCGGAGTCTCTGCT	60
DB	1	ATAAATGACGTGCGAGAGACGAGCGGAGCGGAGCGGAGCGGAGTCTCTGCT	60
QY	61	CCCCCCCCCACCCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	120
DB	61	CCCCCCCCCACCCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	120
QY	121	CACCGCT	180
DB	121	CACCGCT	180
QY	181	TTCCGCT	240
DB	181	TTCCGCT	240
QY	241	CCGAGCTGGAAGCGCCCTCCCGGCTCACTTTTGAACGCTACGCGTCCGCGCG	300
DB	241	CCGAGCTGGAAGCGCCCTCCCGGCTCACTTTTGAACGCTACGCGTCCGCGCG	300
QY	301	TGGAGGTGGGAACAGCGCGCGCATCTCTCCCCCTGGTCAAGCCCAAGCAGCGCG	360

|||||
301 TGGAGTGGGAACAGCGCGCATCTCTCCCTCCCTGGTCAAGCCCAAGCCAGGACGCGCG 360
QY
361 CGGAACCTCTCGGTGTGCTCTCCATCAGTTCGGATCGAGCATCCCCACCAAGCGCT 420
Db
361 CGGAACCTCTCGGTGTGCTCTCCATCAGTTCGGATCGAGCATCCCCACCAAGCGCT 420
QY
421 CACGCCCTCCGGGAGCGCTGGCTTGATACACCGCAGCCCTTCGGGAGACAGCATCTGA 480
Db
421 CACGCCCTCCGGGAGCGCTGGCTTGATACACCGCAGCCCTTCGGGAGACAGCATCTGA 480
QY
481 CTCCCCCGCAGTCAGATTTCCGGACAGCTCTCTAGAAACTCGCTCTTAAGACGGAACCG 540
Db
481 CTCCCCCGCAGTCAGATTTCCGGACAGCTCTCTAGAAACTCGCTCTTAAGACGGAACCG 540
QY
541 CCACAGCACTCAAAAGCCACTCGGAAGAGGAGCGCGGCAAGCCCGGGCCCTGAGCCT 600
Db
541 CCACAGCACTCAAAAGCCACTCGGAAGAGGAGCGCGGCAAGCCCGGGCCCTGAGCCT 600
QY
601 GGACCTTAGCGGTGCCGGCAGCACTGCCGGCGCTTCGCCCTCGCGGACGTCGGCTCCT 660
Db
601 GGACCTTAGCGGTGCCGGCAGCACTGCCGGCGCTTCGCCCTCGCGGACGTCGGCTCCT 660
QY
661 CCTACACTCTCAGCTCGGTGGAGAGACCCCGAGCCCGCCACCATTCAGCGCGGAAGATAC 720
Db
661 CCTACACTCTCAGCTCGGTGGAGAGACCCCGAGCCCGCCACCATTCAGCGCGGAAGATAC 720
QY
721 CTTCCAGATATGCCCTGCTGCTCAAGCCCAATATAGCCTTCCCTCTCAGGTTCCAGTTAT 780
Db
721 CTTCCAGATATGCCCTGCTGCTCAAGCCCAATATAGCCTTCCCTCTCAGGTTCCAGTTAT 780
QY
781 GCGGCGCAGACATACAGCTCGGAATACACCAAGGAGATCATGAACCCCGCATACACCAAG 840
Db
781 GCGGCGCAGACATACAGCTCGGAATACACCAAGGAGATCATGAACCCCGCATACACCAAG 840
QY
841 CTGACATGGAACCTTGGCAGCACTGAGATACAGGCTACAGCCACGACCAAGTCCTGCCAGC 900
Db
841 CTGACATGGAACCTTGGCAGCACTGAGATACAGGCTACAGCCACGACCAAGTCCTGCCAGC 900
QY
901 ATCAGTACTCTGTTGGAGGGTACTCGAGCAACTACGAACCTCAAGCCTTCTCGGTGTAC 960
Db
901 ATCAGTACTCTGTTGGAGGGTACTCGAGCAACTACGAACCTTCTCGGTGTAC 960
QY
961 CAATGCAAGCGCCCTTGAATCAAGTGGAGAGGGGGCGGCGCCAGCTACCATCACCAT 1020
Db
961 CAATGCAAGCGCCCTTGAATCAAGTGGAGAGGGGGCGGCGCCAGCTACCATCACCAT 1020
QY
1021 CACCAACCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 1080
Db
1021 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 1080
QY
1081 CTTCCAGCTCCAGCCGAGGAGAGGTGCTGCCAGCACTCCATGTACTTCAAGCAG 1140
Db
1081 CTTCCAGCTCCAGCCGAGGAGAGGTGCTGCCAGCACTCCATGTACTTCAAGCAG 1140
QY
1141 TCCCCACCGTCAACCCCAACCAACCGGCTTCCCGCGAGGGGGCGGCTTATGGAC 1200
Db
1141 TCCCCACCGTCAACCCCAACCAACCGGCTTCCCGCGAGGGGGCGGCTTATGGAC 1200
QY
1201 GAGGCACTGCCCTCGGCGCCGCTGCTATCGACCCGCGCGCTGCTGGACCCCGCGATG 1260
Db
1201 GAGGCACTGCCCTCGGCGCCGCTGCTATCGACCCGCGCGCTGCTGGACCCCGCGATG 1260
QY
1261 AAGCGGTTCACAGTGGCGCGCGCTTCCGCTTCTTCACTTCAAGCCCTCGCGG 1320
Db
1261 AAGCGGTTCACAGTGGCGCGCGCTTCCGCTTCTTCACTTCAAGCCCTCGCGG 1320
QY
1321 CCGCATCCCCCGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db
1321 CCGCATCCCCCGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY
1381 GCTGCGCGCTCAGGCTGCGCTGGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440

Db
1381 GCTGCGCGCTCAGGCTGCGCTGGGAGCCGAGCCGCCCGCGGCGAGCCAGGCGCGCGCG 1440
QY
1441 CTTGAGAGCCACCCGTACCGGCTGCCCTGCCCAAGAGGGCGGCGCGCTTCCCG 1500
Db
1441 CTTGAGAGCCACCCGTACCGGCTGCCCTGCCCAAGAGGGCGGCGCGCTTCCCG 1500
QY
1501 CTTCTCGGCTCAGCGCCCTCCCTACCGCTCCAGCTGCTGGCGAGAGTCCAGCGCTG 1560
Db
1501 CTTCTCGGCTCAGCGCCCTCCCTACCGCTCCAGCTGCTGGCGAGAGTCCAGCGCTG 1560
QY
1561 CCGTCCGCGCCAGCAGGAGCTCGTCTGCGGAGGCGACGCTGCGCTGCGGCGGAC 1620
Db
1561 CCGTCCGCGCCAGCAGGAGCTCGTCTGCGGAGGCGACGCTGCGCTGCGGCGGAC 1620
QY
1621 AAGCCCGCTCCAGCACTACGCGCTCGAACTTCGAGGGCTCGAAGGGCTTTTCAAG 1680
Db
1621 AAGCCCGCTCCAGCACTACGCGCTCGAACTTCGAGGGCTCGAAGGGCTTTTCAAG 1680
QY
1681 AGAAGTGCAGAAAAATGCAAAATATGTTTGGCTGGCAATAAAACTGCCAGTAGAC 1740
Db
1681 AGAAGTGCAGAAAAATGCAAAATATGTTTGGCTGGCAATAAAACTGCCAGTAGAC 1740
QY
1741 AAGAGAGCTGCAAAACCGATGTCAGTACTGCTGATTTTTCAGAGTCTCTCAGTGTGGAAATG 1800
Db
1741 AAGAGAGCTGCAAAACCGATGTCAGTACTGCTGATTTTTCAGAGTCTCTCAGTGTGGAAATG 1800
QY
1801 GTAAAGAGCTGTCGTCAGATAGTCTGAAAGGAGGAGAGGTCTGCTGCCCTTCAAA 1860
Db
1801 GTAAAGAGCTGTCGTCAGATAGTCTGAAAGGAGGAGAGGTCTGCTGCCCTTCAAA 1860
QY
1861 CCAAGAGCCATTTAACAAGAACCTTCTCAGGCCCTCTCCACCTTCTCTCCTCAATCTGC 1920
Db
1861 CCAAGAGCCATTTAACAAGAACCTTCTCAGGCCCTCTCCACCTTCTCTCCTCAATCTGC 1920
QY
1921 ATGATGATGCTGCTGTCGAGCTTTAAACAGCTTCAACAGCTTCAACAGCTTCTGATTTATTC 1980
Db
1921 ATGATGATGCTGCTGTCGAGCTTTAAACAGCTTCAACAGCTTCAACAGCTTCTGATTTATTC 1980
QY
1981 AGATGATGCTGCTGTCGAGCTTCAACAGCTTCAACAGCTTCAACAGCTTCTGATTTATTC 2040
Db
1981 AGATGATGCTGCTGTCGAGCTTCAACAGCTTCAACAGCTTCAACAGCTTCTGATTTATTC 2040
QY
2041 TACAACTCTGACAGCTTCAATGATGATTCAGAGCTTGGGAGGAGGAGGAGGAGGAGGAG 2100
Db
2041 TACAACTCTGACAGCTTCAATGATGATTCAGAGCTTGGGAGGAGGAGGAGGAGGAGGAG 2100
QY
2101 TTTTACTGATCTCCCAAGAGAGTACAGATTTACTTATGAAATCAGCGCTTTTTCGAGCTG 2160
Db
2101 TTTTACTGATCTCCCAAGAGAGTACAGATTTACTTATGAAATCAGCGCTTTTTCGAGCTG 2160
QY
2161 TTTTCTCTCAGACTTTTCCATCAGGTCAAACTGCTGAAAGATAGTTTGTGTTCTGCAAT 2220
Db
2161 TTTTCTCTCAGACTTTTCCATCAGGTCAAACTGCTGAAAGATAGTTTGTGTTCTGCAAT 2220
QY
2221 GGAATTGCTGATTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db
2221 GGAATTGCTGATTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
QY
2281 AAGACTTTTCTTAAATTTGACAGGCTGAACTTCTGATATCCAAAGCTTTCAGCTGCTG 2340
Db
2281 AAGACTTTTCTTAAATTTGACAGGCTGAACTTCTGATATCCAAAGCTTTCAGCTGCTG 2340
QY
2341 TCAGCACTGAGCATGATCAGAAAGACATGGGTTAAAGAACCAAGAGAGTCCGAAGAG 2400
Db
2341 TCAGCACTGAGCATGATCAGAAAGACATGGGTTAAAGAACCAAGAGAGTCCGAAGAG 2400
QY
2401 CTATGAAAGAGTCAAGAGCTTTAAAGACCAACAGAGTAAAGGAGAGGCTCTGGAG 2460
Db
2401 CTATGAAAGAGTCAAGAGCTTTAAAGACCAACAGAGTAAAGGAGAGGCTCTGGAG 2460
QY
2461 CCCACCGAGTCCAGGCTGCTGGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
Db
2461 CCCACCGAGTCCAGGCTGCTGGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520

```
Qy 2521 CTCAGGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCAATTGAC 2580
Db 2521 CTCAGGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCAATTGAC 2580
Qy 2581 AAGCTCTTCTCGGACACCCCTACTTCTTAATCAGGAGCAGTGGAGCAGTGAAGTGCCTCC 2640
Db 2581 AAGCTCTTCTCGGACACCCCTACTTCTTAATCAGGAGCAGTGGAGCAGTGAAGTGCCTCC 2640
Qy 2641 TCTCTAGCAGCTGTCTGCTACGAGCAAGAGGATAGGTTTGGAAACCTATCATTTCTGTG 2700
Db 2641 TCTCTAGCAGCTGTCTGCTACGAGCAAGAGGATAGGTTTGGAAACCTATCATTTCTGTG 2700
Qy 2701 TCCTTCTTAAAGAGGAAGAGCTCTCTGTAGAAAGCAAGACTTCTTTTTTCTTCTGGC 2760
Db 2701 TCCTTCTTAAAGAGGAAGAGCTCTCTGTAGAAAGCAAGACTTCTTTTTTCTTCTGGC 2760
Qy 2761 TCTTTTCTTCAACACCTAAAGCCAGAAAACTTGACAGATATTTGTTGGGGTGTGTGTTT 2820
Db 2761 TCTTTTCTTCAACACCTAAAGCCAGAAAACTTGACAGATATTTGTTGGGGTGTGTGTTT 2820
Qy 2821 ATATTAGGCATTTGGGGATGGGGTGGAGGGGTTATAGTTCAATGAGGGTTTTCTAAGA 2880
Db 2821 ATATTAGGCATTTGGGGATGGGGTGGAGGGGTTATAGTTCAATGAGGGTTTTCTAAGA 2880
Qy 2881 AATTGCTAAAGAGCACTTTTGGCAATGCTATCCAGCAGCAAAAAAAGGATAATATA 2940
Db 2881 AATTGCTAAAGAGCACTTTTGGCAATGCTATCCAGCAGCAAAAAAAGGATAATATA 2940
Qy 2941 ACTGTTTTAAACCTTTCTCGGGAAATCCAATTATAGTTGCTTTGTTATTTAAACCAAGA 3000
Db 2941 ACTGTTTTAAACCTTTCTCGGGAAATCCAATTATAGTTGCTTTGTTATTTAAACCAAGA 3000
Qy 3001 ACAGCCAAGGGTGTTCGCCAGGCTAGGATGTCTTAAAGATTGGTCCCTTGAATAAT 3060
Db 3001 ACAGCCAAGGGTGTTCGCCAGGCTAGGATGTCTTAAAGATTGGTCCCTTGAATAAT 3060
Qy 3061 GCTTCTGTATCAAGGTACGTATGTGTGCAACAGGCAAGAACTTCTTTTAAATTC 3120
Db 3061 GCTTCTGTATCAAGGTACGTATGTGTGCAACAGGCAAGAACTTCTTTTAAATTC 3120
Qy 3121 CTCTCTCTTTATTTTAAACAATGTTGAAAGATGAGGATTAACCTACAAATCAGACATGG 3180
Db 3121 CTCTCTCTTTATTTTAAACAATGTTGAAAGATGAGGATTAACCTACAAATCAGACATGG 3180
Qy 3181 CAAAACAATAATGCTGTGTCTTCCATAAACAAGTGCATTTTTTAAAGTGTCTCTTA 3240
Db 3181 CAAAACAATAATGCTGTGTCTTCCATAAACAAGTGCATTTTTTAAAGTGTCTCTTA 3240
Qy 3241 CTAAGTCTGTGTTTATTAACCTCTCTTTATTTCTATATGGAATAAAGAGGCGAGTCAATG 3300
Db 3241 CTAAGTCTGTGTTTATTAACCTCTCTTTATTTCTATATGGAATAAAGAGGCGAGTCAATG 3300
Qy 3301 TTAGCAATGACGTTAATAATCCTACGACAGGCTGTGTTTCACTTCCCTGTGCGATCCC 3360
Db 3301 TTAGCAATGACGTTAATAATCCTACGACAGGCTGTGTTTCACTTCCCTGTGCGATCCC 3360
Qy 3361 TTCTGAGGTATGGCCCATCCAAGACTTTTAGGCCATTTCTGATGGAACAGATCCCTGCC 3420
Db 3361 TTCTGAGGTATGGCCCATCCAAGACTTTTAGGCCATTTCTGATGGAACAGATCCCTGCC 3420
Qy 3421 CTGACTGTCCAGCTATCTCTGAAGTGGATCAGATTATAAATGGAATACATGTAATCTGTT 3480
Db 3421 CTGACTGTCCAGCTATCTCTGAAGTGGATCAGATTATAAATGGAATACATGTAATCTGTT 3480
Qy 3481 TTGGTGTGTCTATCAACCCCAACAGAGTCCCTTAACTTGCATTTAGTATAGTAACTG 3540
Db 3481 TTGGTGTGTCTATCAACCCCAACAGAGTCCCTTAACTTGCATTTAGTATAGTAACTG 3540
Qy 3541 ACTGGTATATTCAITTCAGAGCGCCATAAGTCAAGTGTAGTATTTGATCCCTAGATAAGAA 3600
Db 3541 ACTGGTATATTCAITTCAGAGCGCCATAAGTCAAGTGTAGTATTTGATCCCTAGATAAGAA 3600
```

```
Qy 3601 CATCAAAATCAGCAGGAACCTGCTCATACAGGTTAAGCACCCAGGACATTAAGGATTTTTA 3660
Db 3601 CATCAAAATCAGCAGGAACCTGCTCATACAGGTTAAGCACCCAGGACATTAAGGATTTTTA 3660
Qy 3661 TAGATATAAATTAATTTTGTATTGTTAAGGAGACAAATTTTGGAGAGCAAGCAAA--- 3717
Db 3661 TAGATATAAATTAATTTTGTATTGTTAAGGAGACAAATTTTGGAGAGCAAGCAAAATCT 3720
Qy 3718 TCTTTTAAAAAATAGTATGTAATGTGAATCTAGAAAAAGATTTAAAAATAGTATGAGTG 3777
Db 3721 TCTTTTAAAAAATAGTATGTAATGTGAATCTAGAAAAAGATTTAAGAAATAGTATGAGTG 3780
Qy 3778 TGAGTACTAGGAAGAT 3794
Db 3781 TGAGTACTAGGAAGAT 3797

RESULT 2
US-09-949-016-905
; Sequence 905, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 905
; LENGTH: 3802
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-905
```

```
Query Match 98.1%; Score 3722; DB 4; Length 3802;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

Qy 1 ATAATGACGTGCGGAGAGCAGCGAAGCGCAGCGCGGAGCGGAGTCTCTGCT 60
Db 1 ATAATGACGTGCGGAGAGCAGCGAAGCGCAGCGCGGAGCGGAGTCTCTGCT 60
Qy 61 CCGCGCCCGCCACCCCTCCAGCTCTGCTCTCTCCGCTCCCATACACAGCGCGTCA 120
Db 61 CCGCGCCCGCCACCCCTCCAGCTCTGCTCTCTCCGCTCCCATACACAGCGCGTCA 120
Qy 121 CACCGCTCCCTCACTCGCACACACAGCAAGCGGCGCACACAGGCTCCG--CACAC 178
Db 121 CACCGCTCCCTCACTCGAACACACAGCAAGCGGCGCACACAGGCTCCGCGCACAC 180
Qy 179 ACTTCGCTCTCCCGGGGCTCACACCCCTCTTGCCCTGAGCCCTTGGCGGTGACGCGG 238
Db 181 ACTTCGCTCTCCCGGGGCTCACACCCCTCTTGCCCTGAGCCCTTGGCGGTGACGCGG 240
Qy 239 CGCGCAGCTGGACCGCCCTCCCGGGTCACTTTGCAACGCTGAGCGTCCGCGGAGTGGC 298
Db 241 CGCGCAGCTGGACCGCCCTCCCGGGTCACTTTGCAACGCTGAGCGTCCGCGGAGTGGC 300
Qy 299 CGTGGAGTGGGAACAGCGCGGCGATCTCTCCCTCTGCTCAGCCCAAGCCAGACGCC 358
Db 301 CGTGGAGTGGGAACAGCGCGGCGATCTCTCCCTCTGCTCAGCCCAAGCCAGACGCC 360
Qy 359 CGCGGAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATCGCAGCATCCCCCAGCGG 418
Db 361 CGCGGAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATCGCAGCATCCCCCAGCGG 420
```

Qy	419	CTCACGCGCTCCGGGAGCGCTGGGCTTGTACACCGAGCGCTTCCGGGACAGAGCTGT	478
Db	421	CTCACCGCTCCGGGAGCGCTGGGCTTGTACACCGAGCGCTTCCGGGACAGAGCTGT	480
Qy	479	GACTCCCCCAGTGCAGATTTCCGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC	538
Db	481	GACTCCCCCAGTGCAGATTTCCGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC	540
Qy	539	CGCCACAGCACTCAAGCCCACTCGGAAGAGGCGAGCCGGCAAGCCCGGCGCTCAGC	598
Db	541	CGCCACAGCACTCAAGCCCACTCGGAAGAGGCGAGCCGGCAAGCCCGGCGCTCAGC	600
Qy	599	CTGACCCCTTACGCGTTCGGGCGACACTGCGCGCGCTTCCGCTTCGCGGACGCTCCGCTC	658
Db	601	CTGACCCCTTACGCGTTCGGGCGACACTGCGCGCGCTTCCGCTTCGCGGACGCTCCGCTC	660
Qy	659	CTCTACACTCTACGCTTCGCTCGAGAGACCCCGAGCCCACTTACAGCGCGCAAGAT	718
Db	661	CTCTACACTCTACGCTTCGCTCGAGAGACCCCGAGCCCACTTACAGCGCGCAAGAT	720
Qy	719	ACCTCCAGATATGCCCTCGCTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT	778
Db	721	ACCTCCAGATATGCCCTCGCTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT	780
Qy	779	ATGGGGCGCAGACATACAGCTCGGAATACACACCGAGATCATGAACCCCGACTACACCA	838
Db	781	ATGGGGCGCAGACATACAGCTCGGAATACACACCGAGATCATGAACCCCGACTACACCA	840
Qy	839	AGCTGACCATGGACCTTGGGAGGCACTGAGATCAGGCTACAGCACACGCTCCCTGGCCA	898
Db	841	AGCTGACCATGGACCTTGGGAGGCACTGAGATCAGGCTACAGCACACGCTCCCTGGCCA	900
Qy	899	GCATCAGTACCTTCGTGAGGGCTACTCGAGCAACTACGAACTCAAGCCTTCTCGCGTGT	958
Db	901	GCATCAGTACCTTCGTGAGGGCTACTCGAGCAACTACGAACTCAAGCCTTCTCGCGTGT	960
Qy	959	ACCAATGACGCGCCCTTGTAAAGTGGAGGAGGGCGGGCGCCAGCTTACCATCACCC	1018
Db	961	ACCAATGACGCGCCCTTGTAAAGTGGAGGAGGGCGGGCGCCAGCTTACCATCACCC	1020
Qy	1019	ATCAC	1078
Db	1021	ATCAC	1080
Qy	1079	TTCTCTCAGCTCCAGCCCGGAGGACAGAGTGTGCGCCAGCACCTCCCATGTACTTCAAGC	1138
Db	1081	TTCTCTCAGCTCCAGCCCGGAGGACAGAGTGTGCGCCAGCACCTCCCATGTACTTCAAGC	1140
Qy	1139	AGTCCCCACCGTCCACCCCCACACGCGCGCTTCCCGCCAGCGCGGGCGGTATGGG	1198
Db	1141	AGTCCCCACCGTCCACCCCCACACGCGCGCTTCCCGCCAGCGCGGGCGGTATGGG	1200
Qy	1199	ACGAGGCACTGCCCTCGGCGCCCGGCTGCATCGCACCCGCGCCGCTGCTGACCCGCGGA	1258
Db	1201	ACGAGGCACTGCCCTCGGCGCCCGGCTGCATCGCACCCGCGCCGCTGCTGACCCGCGGA	1260
Qy	1259	TGAAGCGCGTCCCAACGCTGGCGGCGCGCTTCCCGCTCTTCCACTTCAAGCCCTCGC	1318
Db	1261	TGAAGCGCGTCCCAACGCTGGCGGCGCGCTTCCCGCTCTTCCACTTCAAGCCCTCGC	1320
Qy	1319	CGCGCATCCCCCGCGCCAGCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1378
Db	1321	CGCGCATCCCCCGCGCCAGCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1380
Qy	1379	CGGCTCCGCGCTCAGCTCGCGTGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1438
Db	1381	CGGCTCCGCGCTCAGCTCGCGTGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1440
Qy	1439	CGCTTGAGAGCACCCGCTACGCGTTCGCGCTGGCCAGAGGCGCGCGCGCGCTTCC	1498
Db	1441	CGCTTGAGAGCACCCGCTACGCGTTCGCGCTGGCCAGAGGCGCGCGCGCGCTTCC	1500

Qy	1499	CGCTCTCGGCTCACGCGCTCCCTTACCGCGTCCAGCTGTCTGGGCGAGAGTCCCGAGCC	1558
Db	1501	CGCTCTCGGCTCACGCGCTCCCTTACCGCGTCCAGCTGTCTGGGCGAGAGTCCCGAGCC	1560
Qy	1559	TGCGGTCCGCGCCAGCAGAGCTCGTGTCTGCGAGGCGCAGTGTGCGGTGCGCGGG	1618
Db	1561	TGCGGTCCGCGCCAGCAGAGCTCGTGTCTGCGAGGCGCAGTGTGCGGTGCGCGGG	1620
Qy	1619	ACAAACGCGCTCCAGCACTACGCGCTGCGAACTCTGCGAGGCTGCAAGGCGCTTTTCA	1678
Db	1621	ACAAACGCGCTCCAGCACTACGCGCTGCGAACTCTGCGAGGCTGCAAGGCGCTTTTCA	1680
Qy	1679	AGAGAACAGTGCAGAAAAATGCAAAATATGTTGCTGCGCAATAAATAATGCGCCAGTAG	1738
Db	1681	AGAGAACAGTGCAGAAAAATGCAAAATATGTTGCTGCGCAATAAATAATGCGCCAGTAG	1740
Qy	1739	ACAAGAGACCTCGAAACCGGATGTCAGTGTGATTTCAAGAGTGTCTCAGTGTTCGAA	1798
Db	1741	ACAAGAGACCTCGAAACCGGATGTCAGTGTGATTTCAAGAGTGTCTCAGTGTTCGAA	1800
Qy	1799	TGTTAAAGAAAGTTGTCGCTACAGATAGTCTGAAAGGGAGAGAGTGTCTGCTTCCCA	1858
Db	1801	TGTTAAAGAAAGTTGTCGCTACAGATAGTCTGAAAGGGAGAGAGTGTCTGCTTCCCA	1860
Qy	1859	AACCAAGAGCCATTAACAACAGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAATCT	1918
Db	1861	AACCAAGAGCCATTAACAACAGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAATCT	1920
Qy	1919	GCATGATGAATGCCCTTGTCCGAGCTTAAACAGACTCAACACCCAGAGATCTTGATTTT	1978
Db	1921	GCATGATGAATGCCCTTGTCCGAGCTTAAACAGACTCAACACCCAGAGATCTTGATTTT	1980
Qy	1979	CCAGATCTGTCCCACTGACAGCTGCTCAGSCACAGATGTGAGCATGTGCAACAAT	2038
Db	1981	CCAGATCTGTCCCACTGACAGCTGCTCAGSCACAGATGTGAGCATGTGCAACAAT	2040
Qy	2039	TCTACAACTCTGACAGCTCCATTTGATGTATCCAGAACTGGGCAAGAAAGTTCGG	2098
Db	2041	TCTACAACTCTGACAGCTCCATTTGATGTATCCAGAACTGGGCAAGAAAGTTCGG	2100
Qy	2099	GATTTACTGATCTCCCAAGAGATCAGACATTTACTTTGATCAGACCTTTTGGAGC	2158
Db	2101	GATTTACTGATCTCCCAAGAGATCAGACATTTACTTTGATCAGACCTTTTGGAGC	2160
Qy	2159	TGTTGTCTCAGACTTTCCATCAGGTCAAACACTGCTGAAGATAAGTTGTGTCTGCA	2218
Db	2161	TGTTGTCTCAGACTTTCCATCAGGTCAAACACTGCTGAAGATAAGTTGTGTCTGCA	2220
Qy	2219	ATGGACTGTCTGCATCGACTTCAGTGCCTTCTGCGATTTTGGGAGTGGCTCGACTCTA	2278
Db	2221	ATGGACTGTCTGCATCGACTTCAGTGCCTTCTGCGATTTTGGGAGTGGCTCGACTCTA	2280
Qy	2279	TTAAAGACTTTTCTTAAATTTGACAGACCTTGAACTTTGATATCCAAAGCTTACCTGCC	2338
Db	2281	TTAAAGACTTTTCTTAAATTTGACAGACCTTGAACTTTGATATCCAAAGCTTACCTGCC	2340
Qy	2339	TGTCAGCACTGAGCATGATCAGAGAAAGATGCGGTTAAAGAAACCAAGAGAGTCCGAG	2398
Db	2341	TGTCAGCACTGAGCATGATCAGAGAAAGATGCGGTTAAAGAAACCAAGAGAGTCCGAG	2400
Qy	2399	AGCTATGCAACAGATCACAAGCAGTTTAAAGAACCCAGAGTAAGGACAGGCTCTGG	2458
Db	2401	AGCTATGCAACAGATCACAAGCAGTTTAAAGAACCCAGAGTAAGGACAGGCTCTGG	2460
Qy	2459	AGCCCAACGAGTCCAAAGTCTCGGTCCTTGGTAGAACTGAGGAGAACTGCAACCTTGG	2518
Db	2461	AACCCAAAGAGTCCAAAGTCTCGGTCCTTGGTAGAACTGAGGAGATCTGCAACCTTGG	2520
Qy	2519	GCCTCCAGCGCATCTTCTAAGCTGAGAGATTTGGTGTCTCCACCTTCCATCATTTG	2578
Db	2521	GCCTCCAGCGCATCTTCTAAGCTGAGAGATTTGGTGTCTCCACCTTCCATCATTTG	2580
Qy	2579	ACAAAGCTTCTCTGGACACACCTTCTTCTTAATCAGGAGCAGTGGAGCAGTGTGCT	2638

Db ||||| 139 CGCCCTCCCGGCTCACTTTGCAACGTGACGGTGGCGGCTGGAGGTGGGA 198
Qy ||||| 312 ACAGGGGGGATCCTCCCGCTGGTCAAGCCCAAGCAGGAGCCCGCGGACCTCTC 371
Db ||||| 199 ACAGGGGGGATCCTCCCGCTGGTCAAGCCCAAGCAGGAGCCCGCGGACCTCTC 258
Qy ||||| 372 GGCTGTCTCTCCCATGAGTGGGATCGCAGCATCCCCCAAGCCGCTCACCCGCTCCG 431
Db ||||| 259 GGCTGTCTCTCCCATGAGTGGGATCGCAGCATCCCCCAAGCCGCTCACCCGCTCCG 318
Qy ||||| 432 GGAGCGCTGGCTTGTACACCGAGCCCTTCGGGAGCAGAGCTGTGAATCCCCCGCAG 491
Db ||||| 319 GGAGCGCTGGCTTGTACACCGAGCCCTTCGGGAGCAGAGCTGTGAATCCCCCGCAG 378
Qy ||||| 492 TGCAGATTTCCGAGCAGCTCTAGAACTCGCTCTAAAGACGGAACCGCCACAGCACTC 551
Db ||||| 379 TGCAGATTTCCGAGCAGCTCTAGAACTCGCTCTAAAGACGGAACCGCCACAGCACTC 438
Qy ||||| 552 AAAGCCACTCGGAAGAGGCGAGCCCGGCAAGCCCGGCCCTGAGCTGGACCTTAGC 611
Db ||||| 439 AAAGCCACTCGGAAGAGGCGAGCCCGGCAAGCCCGGCCCTGAGCTGGACCTTAGC 498
Qy ||||| 612 GGTGGGGGAGCACTGCGCGGCTTTCGCTTCGCGGAGCGTTCGCTCTCACTCTC 671
Db ||||| 499 GGTGGGGGAGCACTGCGCGGCTTTCGCTTCGCGGAGCGTTCGCTCTCACTCTC 558
Qy ||||| 672 AGCTTCGCTGGAGAGACCCCGCCACCATTCAGCGGCAAGATACCTCCAGATAT 731
Db ||||| 559 AGCTTCGCTGGAGAGACCCCGCCACCATTCAGCGGCAAGATACCTCCAGATAT 618
Qy ||||| 732 GCCCTGCGTCAAGCCCAATATAGCCCTTCCAGGTTCCAGTTATGCGGCGAGAC 791
Db ||||| 619 GCCCTGCGTCAAGCCCAATATAGCCCTTCCAGGTTCCAGTTATGCGGCGAGAC 678
Qy ||||| 792 ATAGCTCGGAATACACCAAGAGATCATGAAACCCGAGCTACACCAAGCTGACATGGA 851
Db ||||| 679 ATAGCTCGGAATACACCAAGAGATCATGAAACCCGAGCTACACCAAGCTGACATGGA 738
Qy ||||| 852 CTTGGGAGCACTGAGATCAGGCTACAGCCACACAGCTCCCTGCGCCAGCATCAGTCTT 911
Db ||||| 739 CTTGGGAGCACTGAGATCAGGCTACAGCCACACAGCTCCCTGCGCCAGCATCAGTCTT 798
Qy ||||| 912 CGTGAGGGCTACTCGAGCAACTAGAACTCAAGCTTTCCTGCGTGTACCAAAATGCGAGCG 971
Db ||||| 799 CGTGAGGGCTACTCGAGCAACTAGAACTCAAGCTTTCCTGCGTGTACCAAAATGCGAGCG 858
Qy ||||| 972 GCCCTTGATCAAGTGGAGGAGGGCGGCGCCAGCTACCATCAACCATACACCAACCA 1031
Db ||||| 859 GCCCTTGATCAAGTGGAGGAGGGCGGCGCCAGCTACCATCAACCATACACCAACCA 918
Qy ||||| 1032 CCACCAACCAACCATACAGCAGCAGCATCAGCAGCATCCATTCCTCCAGGCTC 1091
Db ||||| 919 CCACCAACCAACCATACAGCAGCAGCATCAGCAGCATCCATTCCTCCAGGCTC 978
Qy ||||| 1092 CAGCCCGGAGGAGGAGGTGCTGCCAGCACTCCATGTACTTCAAGCAGTCCCCACCGTC 1151
Db ||||| 979 CAGCCCGGAGGAGGAGGTGCTGCCAGCACTCCATGTACTTCAAGCAGTCCCCACCGTC 1038
Qy ||||| 1152 CACCCCAACAGCGCGCTTCCCGCCAGCGGGCGGCGTATGGAGAGGAGCACTGCC 1211
Db ||||| 1039 CACCCCAACAGCGCGCTTCCCGCCAGCGGGCGGCGTATGGAGAGGAGCACTGCC 1098
Qy ||||| 1212 CTCGGCGCCGCTGATCGACCCCGCCGCTGCTGGACCCCGCATGAGAGGCGGTCCC 1271
Db ||||| 1099 CTCGGCGCCGCTGATCGACCCCGCCGCTGCTGGACCCCGCATGAGAGGCGGTCCC 1158
Qy ||||| 1272 CAGGTGGCGGCGCGCTTCCCGCTTTCACATTCAGGCCCTTCGCGCGCATCCCC 1331
Db ||||| 1159 CAGGTGGCGGCGCGCTTCCCGCTTTCACATTCAGGCCCTTCGCGCGCATCCCC 1218
Qy ||||| 1332 CGCGCCAGCGCGCGGCGGCGCACCTCGGCTACGACCGGAGCGGCGCTGGCGGCT 1391

Db ||||| 1219 CGCGCCAGCCCGCGCGCGCCACCACTTCGGGTACGACCCGAGCGCGCGTCCGCGCT 1278
Qy ||||| 1392 CAGCTTCGCGCTGGAGAGCCGCGCGCGCGAGCCAGCGCGCGCGCTTCGAGAGCCA 1451
Db ||||| 1279 CAGCTTCGCGCTGGAGAGCCGCGCGCGCGCGAGCCAGCGCGCGCTTCGAGAGCCA 1338
Qy ||||| 1452 CCCTGAGGGCTCGCTGGCCAAAGAGGGCGCGCGCGCTTCGCGCTTCGCGCT 1511
Db ||||| 1339 CCCTGAGGGCTCGCTGGCCAAAGAGGGCGCGCGCGCTTCGCGCTTCGCGCT 1398
Qy ||||| 1512 CAGCCCTCCCTACCGCTTCAGCGCTTCGCGGAGAGTCCAGCGCTTCGCGCGCC 1571
Db ||||| 1399 CAGCCCTCCCTACCGCTTCGCGGAGAGTCCAGCGCTTCGCGCGCC 1458
Qy ||||| 1572 CAGCAGGAGTCTGCTGCTGGCGAGGGCACTGTGCGCTGTGGGGACAAACCGCGCTG 1631
Db ||||| 1459 CAGCAGGAGTCTGCTGCTGGCGAGGGCACTGTGCGCTGTGGGGACAAACCGCGCTG 1518
Qy ||||| 1632 CAGCAGTACGGCGTGGAAACCTTCGCGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGCA 1691
Db ||||| 1519 CAGCAGTACGGCGTGGAAACCTTCGCGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGCA 1578
Qy ||||| 1692 GAAAGATGCAAAATATGTTTGCCTGGCAATAAACTGCCAGTACCAAGACGCTCG 1751
Db ||||| 1579 GAAAGATGCAAAATATGTTTGCCTGGCAATAAACTGCCAGTACCAAGACGCTCG 1638
Qy ||||| 1752 AAACCGATGTCACTGCTGCTTTCAGAAAGTGTCTCAGTGTGGAAATGGTAAAGAGT 1811
Db ||||| 1639 AAACCGATGTCACTGCTGCTTTCAGAAAGTGTCTCAGTGTGGAAATGGTAAAGAGT 1698
Qy ||||| 1812 TGTCCGTACAGATGTCTGAAAGGGAGAGGTGCTGCTGCTTCCAAACCAAGAGGCC 1871
Db ||||| 1699 TGTCCGTACAGATGTCTGAAAGGGAGAGGTGCTGCTGCTTCCAAACCAAGAGGCC 1758
Qy ||||| 1872 ATTACACAGAGACCTTCAGCCCTTCAGCCCTTCCTCCCAATCTGCAATGATGAATGC 1931
Db ||||| 1759 ATTACACAGAGACCTTCAGCCCTTCAGCCCTTCCTCCCAATCTGCAATGATGAATGC 1818
Qy ||||| 1932 CTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTGATTTTCCAGATACTGCTC 1991
Db ||||| 1819 CTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTGATTTTCCAGATACTGCTC 1878
Qy ||||| 1992 CACTGACAGGCTGCTGAGGCAAGATGTGAGATGTGCAACAATTTACACCTCT 2051
Db ||||| 1879 CACTGACAGGCTGCTGAGGCAAGATGTGAGATGTGCAACAATTTACACCTCT 1938
Qy ||||| 2052 GACAGCTTCCATTTGATGTATCCAGAGCTGGGAGAAAGATTCGGGATTTACTGATCT 2111
Db ||||| 1939 GACAGCTTCCATTTGATGTATCCAGAGCTGGGAGAAAGATTCGGGATTTACTGATCT 1998
Qy ||||| 2112 CCCCAAGAGATCAGACATTTACTTTTGAATCAGCCCTTTTGGAGCTGTTGCTCAG 2171
Db ||||| 1999 CCCCAAGAGATCAGACATTTACTTTTGAATCAGCCCTTTTGGAGCTGTTGCTCAG 2058
Qy ||||| 2172 ACTTTCATCAGGTCAAAACATTCGCTGAAGATTAAGTTTGTGTTCTGCAATGAGACTGCT 2231
Db ||||| 2059 ACTTTCATCAGGTCAAAACATTCGCTGAAGATTAAGTTTGTGTTCTGCAATGAGACTGCT 2118
Qy ||||| 2232 GCATCGACTTCAGTGCCTTCGTTGATTTGGGAGTGGCTCGACTCTATTAAGACTTTTC 2291
Db ||||| 2119 GCATCGACTTCAGTGCCTTCGTTGATTTGGGAGTGGCTCGACTCTATTAAGACTTTTC 2178
Qy ||||| 2292 CTTAAATTTGCGAGGCTGAACCTTTGATATCCAGCCCTTAGCCCTGTGAGCACTGAG 2351
Db ||||| 2179 CTTAAATTTGCGAGGCTGAACCTTTGATATCCAGCCCTTAGCCCTGTGAGCACTGAG 2238
Qy ||||| 2352 CATGATCAAGAGAGATGGGTTAAAGAACCAAGAGAGTCCAGAGCTATGCAACAA 2411
Db ||||| 2239 CATGATCAAGAGAGATGGGTTAAAGAACCAAGAGAGTCCAGAGCTATGCAACAA 2298
Qy ||||| 2412 GATCAAGAGAGTTTAAAGAACCAAGAGTAAAGGAGAGGCTCTGAGGCCACCGAGTC 2471
Db ||||| 2299 GATCAAGAGAGTTTAAAGAACCAAGAGTAAAGGAGAGGCTCTGAGGCCACCGAGTC 2358

Db	586	TTAAGGTAGAAGACATT	CAGATTGCACTACAGCAACACAGCCACTCGTCCCCCAGT	645
Qy	1070	AGCCATCCATTCT	CCAGCTCAGCCCGGAGGACGAGTGTCTGCCAGCACCTCCATGT	1129
Db	646	-----CTGAGGAGATGATG	CGCGCACTCCCGGTGCGTTT	678
Qy	1130	ACTTCAAGAGTCCCA	CGTCCACCCCAACAGCGCGGCTTCCCCTCCAGCGCGGG	1189
Db	679	ACTACAAGCCTCT	CTGCCCCGAGCCACACCCCGGCTTCCAGTGAGCACAGCC	738
Qy	1190	CGTTATGGACGAGCACT	GTCCCTCGGCGCCGGTGTGATCGCACCGCGCCCGCTGCTGG	1249
Db	739	CCATGTGGGACGA	-----CCCGGATCTCTCCACAACTTCCACGAACTACGTGCCA	792
Qy	1250	ACCGCGCGATGAAGCGGT	TCCCAACGGTGGCGCGCGCTTCCGCTCTTCCACTTCA	1309
Db	793	CTACGCACATGATCGAGCAGAGGAAAA	CGCAGTCTCCGCTCTCTCCCTCTTCTCTCTTTA	852
Qy	1310	AGCCTTCGCGCGCAT	CCCCCGCGCCAGCCGCGCGCCACCTCGGCTACG	1369
Db	853	AGCAATGCGCCCCCTGGCA	CCCCGGTGTCTAG-----TT	885
Qy	1370	ACCGCAGCGCGCTGCGCGGCT	CAGCTGCTCGGAGCCGACGCGCGGGGAGCC	1429
Db	886	GCCAGATGCGCTTTCGAGCGGC	CCCTGACGTCCTCCATGAACCCGAGAGCCCGCGCAGCC	945
Qy	1430	AGGCGCGCGCTTGAGAGCCAC	CGGTAGGGCTCGCTGSCCAAGAGGGCGGCCCGC	1489
Db	946	ACCACTGTTGGA	CGGGCAGACTTCGCTGTGCCCAACCCCATTCGCAAGCCGGTCCA	1005
Qy	1490	TGGCTTCCCGCTCTCGGCT	CTCAGCCCTCCCTTACCGCGTCCAGCTGTGGCGAGA	1549
Db	1006	TGGCTTCCCGGCTGCA	GTGGCCACGCGTCTCAGCTGCTGCACGCGAGG-----	1059
Qy	1550	GTCCAGCTTGCCTGCGCGCCAGCAGAGCT	CTGCTCGCTGCGCAGGCGCAGCTGTCGG	1609
Db	1060	-----TGCCCTCACCGCGT	CGCGGGCTCCCTTCAACAGGAGCTGTGCGGTG	1110
Qy	1610	TGTGCGGGACAACGCGCT	CTGCAGCACTACGGGTGCGAACCTGCGAGGGCTCCAAG	1669
Db	1111	TGTGCGGACACGCGCT	CTGCAACATACGGGTGCGCACCTGTGAGGGCTGCAAG	1170
Qy	1670	GCTTTTCAAGAGACA	GTCAGCAAAATATGTTTGCCTGCGCAATAAAACT	1729
Db	1171	GCTTCTTAAAGCAGCAGT	GTCAAAAATATGTTGTTTAGCAATAAAACT	1230
Qy	1730	GCCAGTAGACAAGAGAGCT	CGAAACCGATGTCAGTACTGTGATTTCAAGAGTGTCTCA	1789
Db	1231	GCCAGTGGACAAGCT	TCGCGGNAATCGTGTCAGTACTGCGGATTTCAAGAGTGCCTGG	1290
Qy	1790	GTGTTGGAATGGTAAAA	AGATGTTTCGTTACAGATAGTCTGAAGGAGGAGAGTCTTC	1849
Db	1291	CTGTTGGGATGGTCA	AGAAATGGTTTCGACAGACAGTTTAAAGGCCGAGAGGTCTGT	1350
Qy	1850	TGCCTTCAACACCAAGAGCC	ATTACACAGGAACCTTCTCAGCCCTCTCCACCTTCTC	1909
Db	1351	TGCCCTCGAACCAGAGGCCA	-----CAGGAGCCCTCTCCCCCTTCG	1395
Qy	1910	CTTCAATCTGCATGATGA	TGATGCCCTTTCGAGCTTTAACAGACTCAACCC-----CA	1963
Db	1396	CCCCGGTGGTCTGAT	CAGTGCCCTCGTCAGGGCCCATGTGCACTCAACCCCGGCTATGA	1455
Qy	1964	GAGATCTTGATATTTCAGAT	ACTGTCC-----CACTGACCAAGGCTGTGCAAGGCACAGATG	2020
Db	1456	CCAGCCTCGACTATTTC	CAGGTTCCAGGGGAACCTGACTATCAATCAGTGGAGATGACA	1515
Qy	2021	CTGAGCATGTGC	CAAAATTTCTACAACCTCTCTGACAGCCTCAATGTGATTCAGAGCT	2080
Db	1516	CCCAGCATATCCAGCA	ATTTCTATGATCTCTGACTGCTCATGGATCATCCGGGCT	1575
Qy	2081	GGGCGAAGAAAGAT	TCCGGATTTACTGATCTCCCCCAAGAGATCAACATTACTTATTG	2140
Db	1576	GGGCGAGAGAAGAT	TCCCTGGCTTCGAGACCTTGCCCAAGCGGCAAGACTGCTGTTTG	1635

2321	QY	TCCAAGCCTTAGCCTCGCTGTGACGACTGAGCATGATCA	CAGAAAGACATCGGTTAAAG	2380
1816	Db	TTTCTGCCTTCTCTGTCATTGCTGCCCTATGCTCA	CAGAGAGACACGGGCTCAAG	1875
2381	QY	AACCAAGAGAGTCGAAGACTATGCAACAAGATATCA	CAAGCAGTGTAAAAAGACCAACCAGA	2440
1876	Db	AACCCAAGAGAGTGAAGAAGCTGCAAAAACAGAT	TGTTAAATTTGCTCAAGACCAACGTGA	1935
2441	QY	GTAAGGACAGAGCTCTGGAGCCCAACC-----	GAGTCCAAGTCTCTGGGTGCGCTGG	2491
1936	Db	CTTTCAACAATGGGGGTGGAACGGCCCAATTA	TGTTGTCCTCAAGCTGTGGGAGAGCTCC	1995
2492	QY	TAGAACTGAGGAAGATCTGCACCTCGGCTCCAG	CGCATCTTCTACCTGGAAGCTGGAAG	2551
1996	Db	CAGAACTCGGTACCCCTTGACACAGGGGGCTAC	AGCGCAATTTTCTACCTGGAATTTGGAG	2055
2552	QY	ACTTGGTGTCTCCACCTTCATCAATTGACAAGCT	CTTCTCTGGACACCTTACCTTTCTAA	2610
2056	Db	ACTTGGTGCCACCCGACGCAATATGACAAACT	TTTTTCTCGGACACTTTTACCTTTCTAA	2114
RESULT 7				
US-09-949-016-828				
; Sequence 828, Application US/09949016				
; Patent No. 6812339				
; GENERAL INFORMATION:				
; APPLICANT: VENTER, J. Craig et al.				
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED				
; ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF				
; FILE REFERENCE: CLO01307				
; CURRENT APPLICATION NUMBER: US/09/949,016				
; CURRENT FILING DATE: 2000-04-14				
; PRIOR APPLICATION NUMBER: 60/241,755				
; PRIOR FILING DATE: 2000-10-20				
; PRIOR APPLICATION NUMBER: 60/237,768				
; PRIOR FILING DATE: 2000-10-03				
; PRIOR APPLICATION NUMBER: 60/231,498				
; PRIOR FILING DATE: 2000-09-08				
; NUMBER OF SEQ ID NOS: 207012				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 828				
; LENGTH: 3427				
; TYPE: DNA				
; ORGANISM: Human				
US-09-949-016-828				
Query Match				
Best Local Similarity 13.1%; Score 497; DB 4; Length 3427;				
Matches 1135; Conservative 0; Mismatches 640; Indels 144; Gaps 11;				
QY	722	CTCCAGATATGCCCTGCGTCCAGGCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATG	781	
Db	310	CTGAAGCCATGCTTGTGTTTCAGGCGCAGTAGGGTCTCGCCTCAAGGAGCCAGCCCCG	369	
QY	782	CGGCGCAGACATACAGCTC-----GGAAATACCA	CGGAGATCATGAACCCCG	829
Db	370	CTTCTCAGAGCTACAGTTACCACTCTTCGGGAGAAATACAGTCCGATTTCTTAAC	TCAG	429
QY	830	ACTACACCAAGCTGACCATGGAACCTTGGAGACATGAGATCAGCGGTACAGCCACCA	CGT	889
Db	430	AGTTGTCAAGTTTAGCATGACCTCAACCACTGAAATCACT-----GCCACCACTT	483	
QY	890	CCCTGCCCGACATCAGTACTCTTCGTGGAGGGCTACTCCAGCAA	CTACGAACTCAAGCCTT	949
Db	484	CTCTCCCCAGCTTCAGTACTTTATGGAACACTACAGCACAGGCTACGAGTCAAGCCAC	543	
QY	950	CCTGCGGTGACCAAAATGACGCGCCCTTGATCAAAAGTGGAGAGGGGGCGGCCCA	AGCT	1009
Db	544	CTTGCTTGTAACAAATGCCCCGTGTC-----GGACAGCAGTCTCTCCA	585	
QY	1010	ACCATCACATGACCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	ATCAGCAGCATCAGC	1069


```
Qy 1964 GAGATCTTTGATTATTCAGATACACTGTCC--CACTGACCAGGCTGCTGCAGGCACAGATG 2020
Db 1457 CCAGCTGGACTATTCAGAGTTCCAGGCGAACCTGACTATCAAAATGAGTGGAGATGACA 1516
Qy 2021 CTGAGCATGTGCAACAATTTACAACTCTCCAGAGCTCCATTTGATGTATCCAGAAGCT 2080
Db 1517 CCCAGCATATCCAGCAATTTCTATGATCTCTGACTGCTGCTCCATGAGATCATCCGGGGCT 1576
Qy 2081 GGGCAGAAAAGATTCCGGGATTTACTGATCTCCCAAGAGAGATCAGACATTAATTATG 2140
Db 1577 GGGCAGAGAGATCCCTGGCTTCGAGACCTGCCAAGCGCACAGACCTGTTTTTG 1636
Qy 2141 AATCAGCCTTTTGGAGCTGTTTGTCTCTCAGACTTTCCATCAGGTCAAACTCTCTGAAG 2200
Db 1637 AATCAGCTTTCTTAGAACTGTTTGTCTCTCTGATAGCATACAGTCCAAACCACTGGAGG 1696
Qy 2201 ATAGTTGTGTTCTGCAATGAGCTTGTCTCTGCACTGACTTCACTGCTTCTGTTGATTTG 2260
Db 1697 GTAAACTCATCTTTTGCATGGGTGCTTCTGACAGGTTGCAATGCTTCTGTTGGCTTTG 1756
Qy 2261 GGGAGTGTCTGACTCTATTAAGACTTTTCTTTAAATTTGCAGAGCTGAACCTTGATA 2320
Db 1757 GGGAAATGATTTGATTTCAATGTTGAAATCTCTCACTTCAGATATGACATCGACA 1816
Qy 2321 TCCAAGCCTTAGCCTGCTGTGACACTGAGCATGATCACAAGAAAGATGGGTTAAAG 2380
Db 1817 TTTCTGCTTCTCTGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1876
Qy 2381 AACCAGAGAGTGAAGAGTATGCAACAGATCAGACAGTAAAGACCAACAGA 2440
Db 1877 AACCAGAGAGTGAAGAACTGCAAAACAAAGATTTGTTCTCAAGACCACTGTA 1936
Qy 2441 GTAAGGAGACGCTCTGAGGCCACCC-----GAGTCCAAGTCTGCTGGTGGCTCTGG 2491
Db 1937 CTTTCAACAATGGGGGTTGAAACCGCCCAATTTTGTCCAAACTGTTGGGAAGTCTC 1996
Qy 2492 TAGAACTGAGAAAGATCTGCACTCTGGGCTCCAGCGCATCTTCTACCTGAAGCTGGAAG 2551
Db 1997 CAGAATCTGTAACCTTTGCAACAGGGGCTACAGCGCATTTTCTACCTGAAATTTGGAAG 2056
Qy 2552 ACTTGGTCTCTCCACTTCCATCATTTGACAGCTCTTCTGAGACCCCTACTTCTAA 2610
Db 2057 ACTTGGTCCACCGCCAGCAATAATTGACAAACTTTTCTCTGGACACTTTACCTTTCTAA 2115
```

RESULT 9

```
; Sequence 1, Application US/09277078
; Patent No. 6312949
; GENERAL INFORMATION:
; APPLICANT: Sakurada, Kazuhiro
; APPLICANT: Palmer, Theo
; APPLICANT: Gage, Fred H.
; TITLE OF INVENTION: REGULATION OF TYROSINE HYDROXYLASE
; FILE REFERENCE: 07251/031001
; CURRENT APPLICATION NUMBER: US/09/277,078
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(900)
```

```
Query Match 12.8%; Score 486.4; DB 3; Length 1797;
Best Local Similarity 58.1%; Pred. No. 7,9e-103;
Matches 1105; Conservative 0; Mismatches 671; Indels 126; Gaps 9;
```

```
Qy 730 ATGCCCTTGCTTCAAGCCCAATATAGCCCTTCCCTTCCAGGTTTCCAGTTTANGGGCGCAG 789
Db 1 ATGCTTGTGTTTCAGGCGCAGTATGGTCTCTCGCTCAAGGAGCCAGCCCGCTTCTCAG 60
Qy 790 ACATACAGCT-----CGGATACACCACGAGATCATGAACCCCGACTACACC 837
Db 61 AGTACAGTTTACCACTTCTCGGAGAAATACAGTCCGATTTCTTAACTCCAGAGTTTGTG 120
Qy 838 AAGCTGACATGAGACCTTGGCAGCAGTACAGATCACGGCTACAGCCACCACTGCTCCGCC 897
Db 121 AAGTTAGCATGACCTTCCACCACTGAAT-----TACTGCCACCACTTCTCTCCCC 174
Qy 898 AGCATCAGTACCTTCTGAGGGCTTCTCGAGCAACTACGAACTCAGACCTTCTGCTG 957
Db 175 AGCTTCAGTACCTTTATGGAACAATACAGCACAGGCTTACGACGTCAAGCCACCTTGTG 234
Qy 958 TACCAATGACGGGCCCTTGTATCAAGTGGAGGGGGGGCCCGCCAGCTACCATCAC 1017
Db 235 TACCAATGCCCCCTGTCTC-----GGACAGAGTCTCTCATTTAAGTA 276
Qy 1018 CATCACCAACCAACCAACCAACCACTACCACTACAGAGAGCATCAGAGCCATCC 1077
Db 277 GAAGACATTCAGATGCACAACTACAGCAACACAGCCACCTGCCCTCAGT----- 328
Qy 1078 ATTCTTCCAGCTTCCAGCCCGGAGAGAGGTGTCGCCAGCACTCCATGTTACTTCAAG 1137
Db 329 -----CCGAGGAGATGATGCCACACAGCGGGTCTGGTTTACTACAAG 369
Qy 1138 CAGTCCCCACGTCACCCCAACAGCGCGCTTCCCGCCGAGGGGGGGGTTATGG 1197
Db 370 CCTTTCGGCCCCGACACACCCCGGGTTCAGGTGAGCATAGCCCGATGG 429
Qy 1198 GAGGAGCACTGCTCGGCGCCGCTGCATCGACCCCGCGCTGCTGAGACCCGCCG 1257
Db 430 GACGA-----TCGGGCTCCTTCACTTCCACAGAACTACGTGGCCACTACGCAT 483
Qy 1258 ATGAAGCGGTCCCGTAGGTGGCGCGGCTTCCCGCTCTTCCACTTCAAGCCCTCG 1317
Db 484 ATGATCAGAGCAGAGGAGACACCTGTCTCCCGCTTTCACCTCTTCTTAAAGCAGTCG 543
Qy 1318 CCGCGCATCCCCCGCGCCAGCCCGCGCGGGGCCACCACTCGGCTACGACCCGACG 1377
Db 544 CCGCGGCACTCTCTGTGTAG-----CTGCCAGATG 576
Qy 1378 GCGCTCCCGCTCAGCTTCCGCTGGGAGCCGACCGCGCGGCGAGCAGCGCCGCC 1437
Db 577 CGCTTTGACGGGCTCTGCACGTCCCATGAACCCGAGGCGCGGGCAGCCACAGTA 636
Qy 1438 GCGTTGAGAGCCACCGTAGGGCTCCGCTGGCCAAAGAGGGCGCCCGCTGGCTTC 1497
Db 637 CTGGATGGGAGACCTTTCGCGGTGCCAATCCCAATTCGCAAGCGCGCATCCATGGGCTTC 696
Qy 1498 CCGCTCTCTCGGCTCAGCCCTTCCCTTACCGCGTCCAGCTTGTGGCGAGAGTCCGACG 1557
Db 697 CCGGGCTTCAGATCGGCCAGCGTCTGACAGTGTTCACACGCAAG----- 742
Qy 1558 CTGCGCTCGCGCCACAGAGAGTCTGCTGTGGCGAGGGCAGCTGCTGCGTGTGGGG 1617
Db 743 -TGCCCTCGCGCGCTTCCCGGGCTCTCCCTCCAAATGAGGGTCTGTGCGCTGTGCGGT 801
Qy 1618 GAGAACCGCGCTTCAGACACTAGGGCTGGAACTTCCGAGCTTGTGAGGGCTGCAAGGGCTTTTC 1677
Db 802 GACAAACCGCGCTTCAGCATTTACGGTGTTCGCACTTGTGAGGGCTGCAAGGGTCTTTT 861
Qy 1678 AAGAGAACTGCAAGAAATATGCTTGTCTGCTGCAAAATAAATAAATAAATAAATAAATA 1737
Db 862 AAGCGCAGGTGCAAAACCGCAATATGTGTGTAGCAATAAATAAATAAATAAATAAATA 921
Qy 1738 GACAAGAGAGCTCGAAACCGATGTCACTGTGCTGATTTTCAAGAGTCTCAGTGTGGA 1797
Db 922 GATAAGCGCGCCGAAATTCGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 981
```


2457 GGAGCCACCGAGTCCAAAGTCTCGTGGTCCCTGGTAGAAGTGAAGAGTCTGCACCT 2516
1774 AGCCAGTGCCTGTACAGTCTGTGGGCAAACTCCCGAGCTGGGACCTGTGACCCCA 1833
2517 GGGCTCCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGTGTCTCCACCTTCCATCAT 2576
1834 GGGCTCCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGTGTCTCCACCTTCCATCAT 1893
2577 TGACAAGCTCTTCTGGGACACCTTCTTAATCAGAGCAGTGGAGCAGTGC 2633
1894 TGACAAGATCTTCTATGACAGCTGCTTCTGACCCCTGCTGGGAACAGTGTGC 1950

RESULT 11
US-09-566-921-58
; Sequence 58, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debra W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 58
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 127112.18
US-09-566-921-58

Query Match 9.0%; Score 340.6; DB 4; Length 2637;
Best Local Similarity 61.3%; Pred. No. 7.1e-69;
Matches 663; Conservative 0; Mismatches 389; Indels 29; Gaps 6;

1566 GCGCCCGAGGAGCTGTGTCTGGGAGGACGCTGCGGTGGGGGAGCAAGC 1625
1030 GCGCCCGAGGCGGCGCCAGGTGGAAGTGAAGCCGCTGTGTGTGGGAGCAAGC 1089
1626 CGCTGCGACACTACGCGCTGCAACCTGCGAGGCTGCAAGGCTTTTCAAGAGAAC 1685
1090 TTATGCGACATATGTGTGCGACATGTGAGGCTGCAAGGCTTCTTCAAGCGCAC 1149
1686 AGTGCAGAAAATGCAAAAATATGTTGCTGGCAATATAAATGCGCCAGTAGACAAGAG 1745
1150 AGTGCAGAAAATGCAAAAATATGTTGCTGGCAATATAAATGCGCCAGTAGACAAGAG 1209
1746 AGTGCAGAAAATGCAAAAATATGTTGCTGGCAATATAAATGCGCCAGTAGACAAGAG 1805
1210 GCGCGAAGCGCTGCGACATGTTCCGCTTCCAGAGTGTGCGGTGGGATGTGAA 1269
1806 AGAAGTGTGCGTACAGATAGTGTGAAGGAGGAGAGGTGCTGTGCTTCCAAACCAA 1865
1270 GGAAGTGTGCGACAGACAGCTGGAAGGCGGCGGCGGCTACCTTCAAAACCCAA 1329
1866 GAGCCCATTAACAAGAACCTTCTGAGCCCTTCCACCTTCCATCTGCGATGAT 1925
1330 GCAGCCCCCAGA-----TGCCTCCCTCCCAATCTCTCACTTCCCTGGTCCGT 1378
1926 GAATGCCCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTTATCCAGATA 1985
1379 GGAACACTTGAATCAGGGGCGGCGGCGGCGGCTGCAACTGCAAGTCTCCAGGAGC 1438
1986 CTGTCCCACTGACCCAGGCTGTGAGGAGCAGAGATGTGAGCATGTGCAAAATCTACAA 2045
1439 TGGTCTGCTCCCACTTTTGGGAAGGA-----AGATGCTGGGATGTACAGCAGTCTACGA 1493
2046 CCTCTGACAGCTCCATGATGATATCCAGAAGCTGGGCAAGAAATTCGGGATTTAC 2105

1494 CTTGCTCTCCGTTCTCTGGAGGTCTATCCGCAAGTGGGCGGAGAGATCCCTGGGTTTGC 1553
2106 TGATCTCCCAAGAGATCAGACATTAATTAATCAATCAGCTTTTGGAGCTGTTGT 2165
1554 TGAGCTGTACCGGTGACAGGACCTGTGTGAGTGGGCTTCTTGGAGCTCTTCA 1613
2166 CTTAGAGCTTTC-CATCAGGTCAAAACATGTGTGAAGATAAGTTGTGTCTGCAATGAC 2224
1614 CTTCCGCTGGGTACAAAGTCTAAGCCAGGCGAGGCAAGCTCATCTTCTGCTCAGGCC 1673
2225 TTGCTCTGATC-GACTTCAGTGCCTTCTGAGATTGGGAGTGGCTCGACTCTATTAA 2283
1674 TGGTGTACACCGGCTGCAGTGTGCCCCGTGCGGACTTGGATTGACAGTATCCTG 1733
2284 GACTTTTCTTAATTTTGCAGAGCTTGAACCT-TGATATCCAAAGCTTACCTGCTGCTGC 2342
1734 GCTTCTCAAGTCCCTGCACAGCTTGTGTCGATGTCCCTGCTTGCCTGCTGCTCTC 1793
2343 AGCACTGAGCATGATCAGAAAGACATGGTTTAAAGAACCAAGAGAGTGAAGAGCT 2402
1794 TGCCCTGTCTCATCACCAGCCGATGGCTGTCAGGAGCCGCGGGGTGGAGGCT 1853
2403 ATGCAACAGATCACAAGCAGTTTAAAGAACCAACAGAGTAAGGAGACAGCTCTGGAGCC 2462
1854 GCAGAACCGCATGCCAGCTGCTGAGAGACAGTGGGAGCTGTGGCGGCGAGCC 1913
2463 -----CACCGAGTCCAAGTCTCTGGTGCCTGTGAGTAAGTGAAGAGATCTGCA 2512
1914 AGCCAGCAGCTGCTCTGCTACGCTGTGTTGGGCAAACTGCCGAGCTGCGAGCCTGTGCA 1973
2513 CCCTGGGCTCCAGCGCTCTTCTGAGTGAAGTGAAGACTTGGTGTCTCCACCTTCCA 2572
1974 CCAAGGCTGCGAGCGCATCTTCTACCTCAAGCTGAGAGTGTGGTCCCTCCACCA 2033
2573 TCATTGACAGCTCTTCTGACACACCTTCTTAAATCAGAGAGTGGAGAGTGAAG 2632
2034 TCATTGACAGATCTTCTGACAGCTGCGCTTCTGACCCCTGCTGGGAACACGCTGTG 2093
2633 C 2633
2094 C 2094

RESULT 12
US-09-949-016-35444
; Sequence 35444, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35444
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-35444

Query Match 8.3%; Score 315.6; DB 4; Length 601;
Best Local Similarity 98.8%; Pred. No. 2.2e-63;
Matches 328; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 3466 TTACATGTAACGTTTGGTGTCTATCAACCCACAGGTTCCCTAACTTGCTT 3525
Db 1 TTACATGTAACGTTTGGTGTCTATCAACCCACAGGTTCCCTAACTTGCTT 60
QY 3526 CAGTTATAGTAACGACTGGTATATTCATTCAAGACGCCATAAGTCAGTTGAGTATTG 3585
Db 61 CAGTTATAGTAACGACTGGTATATTCATTCAAGACGCCATAAGTCAGTTGAGTATTG 120
QY 3586 ATCCCTAGATAAGACATCAAAATCAGCAGGAACTGGTCATACAGGGTAAGCACCAGGA 3645
Db 121 ATCCCTAGATAAGACATCAAAATCAGCAGGAACTGGTCATACAGGGTAAGCACCAGGA 180
QY 3646 CAATAAGGATTTTATAGATATAATTTTATTTTGTGTTAAGGAGACAAATTTGG 3705
Db 181 CAATAAGGATTTTATAGATATAATTTTATTTTGTGTTAAGGAGACAAATTTGG 240
QY 3706 AGAGCAAGCAAA--TCCTTTTAAAAAATAGTATGAATGTGAATCTAGAAAAGATTTAA 3762
Db 241 AGAGCAAGCAAAATCTCTTTTAAAAAATAGTATGAATGTGAATCTAGAAAAGATTTAA 300
QY 3763 AAAATAGTATGAGTGTGAGTACTAGGAAGGAT 3794
Db 301 AAAATAGTATGAGTGTGAGTACTAGGAAGGAT 332

RESULT 13

US-09-949-016-135237
; Sequence 135237, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 135237
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-135237

Query Match 8.3%; Score 315.6; DB 4; Length 601;
Best Local Similarity 98.8%; Pred. No. 2.2e-63;
Matches 328; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 3466 TTACATGTAACGTTTGGTGTCTATCAACCCACAGGTTCCCTAACTTGCTT 3525
Db 1 TTACATGTAACGTTTGGTGTCTATCAACCCACAGGTTCCCTAACTTGCTT 60
QY 3526 CAGTTATAGTAACGACTGGTATATTCATTCAAGACGCCATAAGTCAGTTGAGTATTG 3585
Db 61 CAGTTATAGTAACGACTGGTATATTCATTCAAGACGCCATAAGTCAGTTGAGTATTG 120
QY 3586 ATCCCTAGATAAGACATCAAAATCAGCAGGAACTGGTCATACAGGGTAAGCACCAGGA 3645
Db 121 ATCCCTAGATAAGACATCAAAATCAGCAGGAACTGGTCATACAGGGTAAGCACCAGGA 180
QY 3646 CAATAAGGATTTTATAGATATAATTTTATTTTGTGTTAAGGAGACAAATTTGG 3705
Db 181 CAATAAGGATTTTATAGATATAATTTTATTTTGTGTTAAGGAGACAAATTTGG 240
QY 3706 AGAGCAAGCAAA--TCCTTTTAAAAAATAGTATGAATGTGAATCTAGAAAAGATTTAA 3762
Db 241 AGAGCAAGCAAAATCTCTTTTAAAAAATAGTATGAATGTGAATCTAGAAAAGATTTAA 300

QY 3763 AAAATAGTATGAGTGTGAGTACTAGGAAGGAT 3794
Db 301 AAAATAGTATGAGTGTGAGTACTAGGAAGGAT 332

RESULT 14

US-09-949-016-135236
; Sequence 135236, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 135236
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-135236

Query Match 6.2%; Score 236; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 7e-45;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2360 CAGAAGACATGGGTTAAAGAACCAAGAGAGTCAAGAGCTATGCAACAAGATCACAA 2419
Db 366 CAGAAGACATGGGTTAAAGAACCAAGAGAGTCAAGAGCTATGCAACAAGATCACAA 425
QY 2420 GCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCTGGAGCCCAAGGTC 2479
Db 426 GCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCTGGAGCCCAAGGTC 485
QY 2480 TGGTGCCCTGGTAGAATCTGAGGAGATCTGCACCTTCCATCATTTCTCTGGAC 2539
Db 486 TGGTGCCCTGGTAGAATCTGAGGAGATCTGCACCTTCCATCATTTCTCTGGAC 545
QY 2540 TGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTCTCTGGAC 2595
Db 546 TGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTCTCTGGAC 601

RESULT 15

US-09-949-016-35441
; Sequence 35441, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35441
; LENGTH: 601

Tue Mar 22 09:54:21 2005

```
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-35441

Query Match      4.6%; Score 174.6; DB 4; Length 601;
Best Local Similarity 99.4%; Pred. No. 1.3e-30;
Matches 174; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1809 AGTTGTCCGTACAGATAGTCTGAAAGGAGGAGAGGTCGTCTGCTTCCAAACCAAAGAG 1868
Db      |||:|||||
QY 178  AGTTGTCCGTACAGATAGTCTGAAAGGAGGAGAGGTCGTCTGCTTCCAAACCAAAGAG 237
Db      |||:|||||
QY 1869 CCCATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCAAATCTGCATGATGAA 1928
Db      |||:|||||
QY 238  CCCATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCAAATCTGCATGATGAA 297
Db      |||:|||||
QY 1929 TGCCCTTGTCCGAGCTTTTAAACAGACTCAACACCCAGAGATCTTGATTATTCAGA 1983
Db      |||:|||||
QY 298  TGCYCTTGTCCGAGCTTTTAAACAGACTCAACACCCAGAGATCTTGATTATTCAGA 352
Db      |||:|||||
```

Search completed: March 21, 2005, 20:50:42
Job time : 606 secs

Qy	2221	GGACTTGTCTCTGCAATCGACTTCAGTGCCTTCGTGGATTTTGGGAGTGGCTCGACTCTATTT	2280
Db	2221	GGACTTGTCTCTGCAATCGACTTCAGTGCCTTCGTGGATTTTGGGAGTGGCTCGACTCTATTT	2280
Qy	2281	AAAGACTTTTTCCTTAAATTTTCAGAGCCTCGAACCTTCATATCCAAAGCCTTAGCCTGCCTG	2340
Db	2281	AAAGACTTTTTCCTTAAATTTTCAGAGCCTCGAACCTTCATATCCAAAGCCTTAGCCTGCCTG	2340
Qy	2341	TCAGCACTGAGCATGATCACAGAAAGACATGGGTTTAAAGAAACCAAAGAGAGTCGAAGAG	2400
Db	2341	TCAGCACTGAGCATGATCACAGAAAGACATGGGTTTAAAGAAACCAAAGAGAGTCGAAGAG	2400
Qy	2401	CTATGCAACAAGATCACAGCAGTTTAAAAGAACAACAGAGTAAAGGACAGGCTCTGGAG	2460
Db	2401	CTATGCAACAAGATCACAGCAGTTTAAAAGACCAACAGAGTAAAGGACAGGCTCTGGAG	2460
Qy	2461	CCACCGAGTCCAAAGTCTGGGTGCCCTGTGTAGAACTGAGGAGAGTCTGCACCCCTGGGC	2520
Db	2461	CCACCGAGTCCAAAGTCTGGGTGCCCTGTGTAGAACTGAGGAGAGTCTGCACCCCTGGGC	2520
Qy	2521	CTCCAGGCGCATCTTCTACCTGAAGCTGGAAAGACTTGGTGTCTCCACCTTCCATCATTTGAC	2580
Db	2521	CTCCAGGCGCATCTTCTACCTGAAGCTGGAAAGACTTGGTGTCTCCACCTTCCATCATTTGAC	2580
Qy	2581	AAGCTCTTCTGGAACACCTTCTTAATTCAGGAGCAGTGGAGCAGTGCAGTGCCTGCC	2640
Db	2581	AAGCTCTTCTGGAACACCTTCTTAATTCAGGAGCAGTGGAGCAGTGCAGTGCCTGCC	2640
Qy	2641	TCCTCTAGCACTCTGCTACCGCAGCAAAAGGATAGTTTGGAAACCTATCATTTCCCTG	2700
Db	2641	TCCTCTAGCACTCTGCTACCGCAGCAAAAGGATAGTTTGGAAACCTATCATTTCCCTG	2700
Qy	2701	TCCTTCTCTTAAGAGGAAAAGCAGCTCCTCTGTAGAAAGCAAAAGACTTCTTTTTTTTCTGGC	2760
Db	2701	TCCTTCTCTTAAGAGGAAAAGCAGCTCCTGTGTAGAAAGCAAAAGACTTCTTTTTTTTCTGGC	2760
Qy	2761	TCTTTTCTCTTAAACCTTAAAGCCAGAAAACTTGCAGAGTATGTGTGTGGGGTGTGTGTTT	2820
Db	2761	TCTTTTCTCTTAAACCTTAAAGCCAGAAAACTTGCAGAGTATGTGTGTGGGGTGTGTGTTT	2820
Qy	2821	ATATTTAGGCATTTGGGGATGGGTGGAGGGGGTTATAGTTCATGAGGGTTCCTTAAGA	2880
Db	2821	ATATTTAGGCATTTGGGGATGGGTGGAGGGGGTTATAGTTCATGAGGGTTCCTTAAGA	2880
Qy	2881	AAATGCTTAAACAAAGCACTTTTGGACAATGCTATCCACAGGAGGAAAAAAGGATATATA	2940
Db	2881	AAATGCTTAAACAAAGCACTTTTGGACAATGCTATCCACAGGAGGAAAAAAGGATATATA	2940
Qy	2941	ACTGTTTTTAAACTCTTTCTGGGGAATCCAAATATAGTTCCTTGTATTTTAAAAACAAGA	3000
Db	2941	ACTGTTTTTAAACTCTTTCTGGGGAATCCAAATATAGTTCCTTGTATTTTAAAAACAAGA	3000
Qy	3001	ACAGCCAAGGGTTGTCGCCAGGTAGAGTGTCTTTAAAGATTGTTCCTTGAANAATAT	3060
Db	3001	ACAGCCAAGGGTTGTCGCCAGGTAGAGTGTCTTTAAAGATTGTTCCTTGAANAATAT	3060
Qy	3061	GCCTCTCTGATCAAAGGTACGTATGTGGTGCMAACAGGAGGAGAACTTCCTTTTAAATTC	3120
Db	3061	GCCTCTCTGATCAAAGGTACGTATGTGGTGCMAACAGGAGGAGAACTTCCTTTTAAATTC	3120
Qy	3121	CTTCTCTCTTTATTTTAAACAAATGGTGAAGATGGAGGATTAACCTACAAATCAGACATGG	3180
Db	3121	CTTCTCTCTTTATTTTAAACAAATGGTGAAGATGGAGGATTAACCTACAAATCAGACATGG	3180
Qy	3181	CAAAACAATAATGGCTGTTTCTCCATTAACAAGTGCATATTTTAAAGTGTCTGCTTA	3240
Db	3181	CAAAACAATAATGGCTGTTTCTCCATTAACAAGTGCATATTTTAAAGTGTCTGCTTA	3240
Qy	3241	CTAAGTCTGTGTTTATTAACCTCTCCTTTATTTCTATATGGAATAAAAAGGAGGAGTCATG	3300
Db	3241	CTAAGTCTGTGTTTATTAACCTCTCCTTTATTTCTATATGGAATAAAAAGGAGGAGTCATG	3300
Qy	3301	TTAGCAAAATGACAGCTTAAATATCCCTACGAGAGGCTGTGTTCACCTTCCCTGTGCGATCCC	3360

Db	3301	TTAGCAAAATGACACGGTTAAATATCCCTAGCAGAGGCTGTGTTCACCTTCCCTGTGCGATCCC	3360
Qy	3361	TTCTGAGGTATGGCCCATCCAAGACTTTTATAGGCCATTCCTTGATGGAAACCAAGATCCCTGTC	3420
Db	3361	TTCTGAGGTATGGCCCATCCAAGACTTTTATAGGCCATTCCTTGATGGAAACCAAGATCCCTGTC	3420
Qy	3421	CTGACTGTCCAGCTATCCTGAAAGTGGATCAGATTTATAAATCTGGAATTAACATGTAACCTGTT	3480
Db	3421	CTGACTGTCCAGCTATCCTGAAAGTGGATCAGATTTATAAATCTGGAATTTACATGTAACCTGTT	3480
Qy	3481	TTGGTGTGTCTATCAACCCCAAGAGTTCCTTAAACTTGGCTTCAGTTATAGTAACCTG	3540
Db	3481	TTGGTGTGTCTATCAACCCCAAGAGTTCCTTAAACTTGGCTTCAGTTATAGTAACCTG	3540
Qy	3541	ACTGGTATATTCATTCAAGAACGCGCCATAAGTCAGTTGAGTATTTGATCCCTAGATAAGAA	3600
Db	3541	ACTGGTATATTCATTCAAGAACGCGCCATAAGTCAGTTGAGTATTTGATCCCTAGATAAGAA	3600
Qy	3601	CATGCCAAATCAGCAGGAACCTGGTCATACAGGGTAAAGCACACAGGACAATAAGGATTTTTA	3660
Db	3601	CATGCCAAATCAGCAGGAACCTGGTCATACAGGGTAAAGCACACAGGACAATAAGGATTTTTA	3660
Qy	3661	TAGATATAATTTAAATTTTTCGTATTGGTTAAGGAGACAAATTTTGGAGAGCAACCAATCT	3720
Db	3661	TAGATATAATTTAAATTTTTCGTATTGGTTAAGGAGACAAATTTTGGAGAGCAACCAATCT	3720
Qy	3721	TTTTTAAAAAATAGTATGAAATGTGAATCTAGAAAAAGATTTTAAAAAATAGTATGAGTGTGA	3780
Db	3721	TTTTTAAAAAATAGTATGAAATGTGAATCTAGAAAAAGATTTTAAAAAATAGTATGAGTGTGA	3780
Qy	3781	GTACTAGGAAGGAT	3794
Db	3781	GTACTAGGAAGGAT	3794

RESULT 2

US-09-386-46
; Sequence 46, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Bresnihan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
; TITLE OF INVENTION: Subfamily of Nuclear Transcription Factors
; FILE REFERENCE: P01972US1
; CURRENT APPLICATION NUMBER: US/09/853,386
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/203645
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 3802
; TYPE: DNA
; ORGANISM: HUMAN
US-09-853-386-46

[illegible]


```
QY 2339 TGTGACGCTGAGCATGATCAGAGAGACATGGGTTAAAGAACCAAGAGAGTGAAG 2398
Db 2341 TGTGACGCTGAGCATGATCAGAGAGACATGGGTTAAAGAGAACCAAGAGAGTGAAG 2400
QY 2399 AGCTATGCAACAAGATCACAAGCAGTTTAAAGAGACCAACAGAGTAAAGGACAGGCTCTGG 2458
Db 2401 AGCTATGCAACAAGATCACAAGCAGTTTAAAGAGACCAACAGAGTAAAGGACAGGCTCTGG 2460
QY 2459 AGCCACCGAGTCCAAAGTCTCGGTCCTGGTAGAACTGAGAGAGATCTCGACCCCTGG 2518
Db 2461 AACCACCAAGTCCAAAGTCTCGGTCCTGGTAGAACTGAGAGAGATCTCGACCCCTGG 2520
QY 2519 GCTCCAGCGCATCTCTACCTGAAGCTGGAAGCTTGGTGTCTCCACCTTCATCATTTG 2578
Db 2521 GCTCCAGCGCATCTCTACCTGAAGCTGGAAGCTTGGTGTCTCCACCTTCATCATTTG 2580
QY 2579 ACAAGCTCTTCTGACACACCTTCTCTAATCAGGAGCAGTGGAGCAGTGAAGTGCCT 2638
Db 2581 ACAAGCTCTTCTGACACACCTTCTCTAATCAGGAGCAGTGGAGCAGTGAAGTGCCT 2640
QY 2639 CTTCTCTAGCAGCTGCTTGTGACGAGCAAAAGGATAGGTTTGAAGCTATCATTTCC 2698
Db 2641 CTTCTCTAGCAGCTGCTTGTGACGAGCAAAAGGATAGGTTTGAAGCTATCATTTCC 2700
QY 2699 TGTCTCTTCTTAAAGGAAAGCAGCTCTGTAGAAAGCAAGACTTCTTTTCTTTCTG 2758
Db 2701 TGTCTCTTCTTAAAGGAAAGCAGCTCTGTAGAAAGCAAGACTTCTTTTCTTTCTG 2760
QY 2759 GCTCTTTTCTTACAACTTAAAGCCAGAAATCTGCAGAGTATTTGTTGGGTTGTGTT 2818
Db 2761 GCTCTTTTCTTACAACTTAAAGCCAGAAATCTGCAGAGTATTTGTTGGGTTGTGTT 2820
QY 2819 TTATATTAGGCAATGGGGATGGGGTGGAGGGGTTATAGTTTATGAGGGTTTCTAA 2878
Db 2821 TTATATTAGGCAATGGGGATGGGGTGGAGGGGTTATAGTTTATGAGGGTTTCTAA 2880
QY 2879 GAAATGCTTAAAGACCTTTTGGACATGCTATCCAGCAGGAGGAGGAGGAGGATATA 2938
Db 2881 GAAATGCTTAAAGACCTTTTGGACATGCTATCCAGCAGGAGGAGGAGGAGGATATA 2940
QY 2939 TAACTGTTTAAACCTTCTTCTGGGAATCCAAATATAGTCTCTTGTATTTAAAGCAA 2998
Db 2941 TAACTGTTTAAACCTTCTTCTGGGAATCCAAATATAGTCTCTTGTATTTAAAGCAA 3000
QY 2999 GAACAGCCAGGGTTTTCGCCAGGGTAGGATGTGCTTTAAAGATTTGGTCCCTTGAAT 3058
Db 3001 GAACAGCCAGGGTTTTCGCCAGGGTAGGATGTGCTTTAAAGATTTGGTCCCTTGAAT 3060
QY 3059 ATGCTTCTGTATCAAGGTAGTATGTGGTCAACAAAGGAGGAGGAGGAGGATTAAT 3118
Db 3061 ATGCTTCTGTATCAAGGTAGTATGTGGTCAACAAAGGAGGAGGAGGAGGATTAAT 3120
QY 3119 TCCTTCTTCTTTTAAACAAATGGTGAAAGATGGAGGATTAACCTACAAATCAGACAT 3178
Db 3121 TCCTTCTTCTTTTAAACAAATGGTGAAAGATGGAGGATTAACCTACAAATCAGACAT 3180
QY 3179 GCGAAAACAATATGGCTTTTCTTCCATAACCAAGTGCAATTTTTTAAAGTGTGTCT 3238
Db 3181 GCGAAAACAATATGGCTTTTCTTCCATAACCAAGTGCAATTTTTTAAAGTGTGTCT 3240
QY 3239 TACTAAGTCTGTATTAACCTCTCTTTTATTTATATGGAATATAAAGAGGAGGAGTCA 3298
Db 3241 TACTAAGTCTGTATTAACCTCTCTTTTATTTATATGGAATATAAAGAGGAGGAGTCA 3300
QY 3299 TGTATGAAAATCAGACGTTTAAATATCCCTAGCAGAGCTGTGTTTCACTTCCCTGTGATC 3358
Db 3301 TGTATGAAAATCAGACGTTTAAATATCCCTAGCAGAGCTGTGTTTCACTTCCCTGTGATC 3360
QY 3359 CTTCTGAGGATATGGCCCATCAAGACTTTTATAGGCCATTTCTGTATGGAACCAAGATCCCTG 3418
Db 3361 CTTCTGAGGATATGGCCCATCAAGACTTTTATAGGCCATTTCTGTATGGAACCAAGATCCCTG 3420
```

```
QY 3419 CCTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTATATAACTGGATTACATGTAACCTG 3478
Db 3421 CCTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTATATAACTGGATTACATGTAACCTG 3480
QY 3479 TTTTGGTTGTGTTCTTATCAACCCCAACAGAGTTCCCTAAACCTTGTCTTCAAGTTATAGTAAC 3538
Db 3481 TTTTGGTTGTGTTCTTATCAACCCCAACAGAGTTCCCTAAACCTTGTCTTCAAGTTATAGTAAC 3540
QY 3539 TGACTGGTATATTCATTTGAGAGGCCCATTAAGTCAGTTAGTATTTGATCCCTAGATAAG 3598
Db 3541 TGACTGGTATATTCATTTGAGAGGCCCATTAAGTCAGTTAGTATTTGATCCCTAGATAAG 3600
QY 3599 AACATGCAAAATCAGCAGGAACTGGTCTACAGGGTAAGCACCAGGACAAATAGGATTTT 3658
Db 3601 AACATGCAAAATCAGCAGGAACTGGTCTACAGGGTAAGCACCAGGACAAATAGGATTTT 3660
QY 3659 TATAGATATAATTTAAATTTTGTTA- - TTGGTTAAAGGAGA- -CAATTTTGGAGAGCAAGCA 3715
Db 3661 TATAGATATAATTTAAATTTTGTTAATTTGGTAAATTTGGTAAAGGACCAATTTTGGAGAGCAAGCA 3720
QY 3716 AA- - -TCTTTTAAATAATAGTATGAATGTGAATACTAGAAAAGATTTTAAATAATAGTAT 3772
Db 3721 AATCTTCTTTTAAATAATAGTATGAATGTGAATACTAGAAAAGATTTTAAATAATAGTAT 3780
QY 3773 GAGTGTGAGTACTAGGAGGAT 3794
Db 3781 GAGTGTGAGTACTAGGAGGAT 3802
```

RESULT 3

```
US-09-853-386-78
; Sequence 78, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Breenihan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
; FILE REFERENCE: P01972US1
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/203645
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
; LENGTH: 3802
; TYPE: DNA
; ORGANISM: HUMAN
US-09-853-386-78
```

```
Query Match 98.1%; Score 3722; DB 9; Length 3802;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;
```

```
QY 1 ATAAATGAGTGCAGAGAGCGAGCAACGCGAGCGGAGAGCGGAGTCTCTGCT 60
Db 1 ATAAATGAGTGCAGAGAGCGAGCAACGCGAGCGGAGAGCGGAGTCTCTGCT 60
QY 61 CCGGCCCCCAACCCCTCCAGTCTCTGCTCTCTCCGCTCCCAATACACAGCGGCTCA 120
Db 61 CCGGCCCCCAACCCCTCCAGTCTCTGCTCTCTCCGCTCCCAATACACAGCGGCTCA 120
QY 121 CACCGCTCTCCCTCTCACTCCGACACAGAGCGGCGGACACAGGCTCCG--CACAC 178
Db 121 CACCGCTCTCCCTCTCACTCCGACACAGAGCGGCGGACACAGGCTCCGCGACACAC 180
QY 179 ACTTGGCTCTCCCGCGGCTCACACCCCTCTTGGCTTGGCCCTTGGCGGCGG 238
Db 181 ACTTGGCTCTCCCGCGGCTCACACCCCTCTTGGCTTGGCCCTTGGCGGCGG 240
```

QY 239 CGCCGACGTGAGCGCCCTCCGGGCTCACTTTGCAAGCTGACGGTCCGGCAGTGGC 298
DB 241 CGCCGACGTGAGCGCCCTCCGGGCTCACTTTGCAAGCTGACGGTCCGGCAGTGGC 300
QY 299 CGTGAAGTGGGAAACAGCGGGCGGCAATCTCCGCCCTGGTCAAGCCCAAGCCAGACGCC 358
DB 301 CGTGAAGTGGGAAACAGCGGGCGGCAATCTCCGCCCTGGTCAAGCCCAAGCCAGACGCC 360
QY 359 CGCGAAACCTCTCGGCTGCTCCCATAGTTCGGGATCGCAGCATCCCCACCAGCCG 418
DB 361 CGCGAAACCTCTCGGCTGCTCCCATAGTTCGGGATCGCAGCATCCCCACCAGCCG 420
QY 419 CTCACCGCTCCGGAGCGGCTGGCTTTGTACACCGCAGCCCTTCGGGACAGCAGCTGT 478
DB 421 CTCACCGCTCCGGAGCGGCTGGCTTTGTACACCGCAGCCCTTCGGGACAGCAGCTGT 480
QY 479 GACTCCCGCCAGTCAGATTTGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC 538
DB 481 GACTCCCGCCAGTCAGATTTGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC 540
QY 539 CGCCACAGCACTCAAGGCCACTGCGGAAGAGGCGAGCCGCAAGCCCGGCCCTGAGC 598
DB 541 CGCACAAGCACTCAAGGCCACTGCGGAAGAGGCGAGCCGCAAGCCCGGCCCTGAGC 600
QY 599 CTGGAACCTTAGCGGTGCGGGCAGCACTGCGGGCTTCGGCTTCGGCGGAGCTCCGCTC 658
DB 601 CTGGAACCTTAGCGGTGCGGGCAGCACTGCGGGCTTCGGCTTCGGCGGAGCTCCGCTC 660
QY 659 CTCCTACACTCTCAGCTTCGGTGGAGAGACCCCGAGCCCAACATTCAGCGCGCAAGAT 718
DB 661 CTCCTACACTCTCAGCTTCGGTGGAGAGACCCCGAGCCCAACATTCAGCGCGCAAGAT 720
QY 719 ACCCTCCAGATATGCCCTGCGTCCAAAGCCCAATATAGCCCTTCAGCTTCAGTT 778
DB 721 ACCCTCCAGATATGCCCTGCGTCCAAAGCCCAATATAGCCCTTCAGCTTCAGTT 780
QY 779 ATGCGCGCAGACATACAGCTCGGAATACACCGAGATCATGAACCCCGACTACACA 838
DB 781 ATGCGCGCAGACATACAGCTCGGAATACACCGAGATCATGAACCCCGACTACACA 840
QY 839 AGCTGACCATGACCTTGGCAGCACTGAGATCAGGCTACGCCACCAAGCTCCCTGCCA 898
DB 841 AGCTGACCATGACCTTGGCAGCACTGAGATCAGGCTACGCCACCAAGCTCCCTGCCA 900
QY 899 GCATCAGTACCTTCGTGAGGCTACTCGAGCAACTACGAACTCAAGCTTCCTCGCTGT 958
DB 901 GCATCAGTACCTTCGTGAGGCTACTCGAGCAACTACGAACTCAAGCTTCCTCGCTGT 960
QY 959 ACCAAATGACGGCCCTTGATCAAAAGTGGAGAGGGCGGCGGCCAGCTACCATACC 1018
DB 961 ACCAAATGACGGCCCTTGATCAAAAGTGGAGAGGGCGGCGGCCAGCTACCATACC 1020
QY 1019 ATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1078
DB 1021 ATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1080
QY 1079 TTTCTCCAGCTCCAGCCCGAGGACGAGTGTCTGCCAGCACTCCATGACTTCAAGC 1138
DB 1081 TTTCTCCAGCTCCAGCCCGAGGACGAGTGTCTGCCAGCACTCCATGACTTCAAGC 1140
QY 1139 AGTCCCAACCGTCCACCCCAACCAAGTGGAGAGGGCGGCGGCCAGCTACCATACC 1198
DB 1141 AGTCCCAACCGTCCACCCCAACCAAGTGGAGAGGGCGGCGGCCAGCTACCATACC 1200
QY 1199 ACGAGCACTGCTCCGCGCGCGGCTGCATCGACCCGCGCGCTGCGACCCGCGCA 1258
DB 1201 ACGAGCACTGCTCCGCGCGCGGCTGCATCGACCCGCGCGCTGCGACCCGCGCA 1260
QY 1259 TGAAGCGGTCCCAACCGTGGCGCGCGCTTCCTCCAGCTTCAGAGCCCTGCG 1318
DB 1261 TGAAGCGGTCCCAACCGTGGCGCGCGCTTCCTCCAGCTTCAGAGCCCTGCG 1320
QY 1319 CGCCGATCCCCCGGCGCCAGCGCGCGCGGCGCAACACCTCGGCTACGACCCGAGCG 1378

DB 1321 CGCCGATCCCGCGCGCCAGCCCGCGCGCGCCACCACTCGGCTACACCCGACGG 1380
QY 1379 CGCTGCGCGCTCAGCTGCGCTGGAGCGCGCAGCGCGCGCGGACGAGCGCGCG 1438
DB 1381 CGCTGCGCGCTCAGCTGCGCTGGAGCGCGCAGCGCGCGCGGACGAGCGCGCG 1440
QY 1439 CGCTTGAGAGCACCCGCTACCGGCTGCGCTGGCAAGAGGCGCGCGCTGCTTCC 1498
DB 1441 CGCTTGAGGCGCAACCGCTACCGGCTGCGCTGGCCNAGAGGCGCGCGCTGCTTCC 1500
QY 1499 CGCTCTCGGCTCAGCGCTCCCTACCGCTCAGCTGCGCTGGGCGAGAGTCCAGCC 1558
DB 1501 CGCTCTCGGCTCAGCGCTCCCTACCGCTCAGCTGCGCTGGGCGAGAGTCCAGCC 1560
QY 1559 TGCCGTGCGCGCCAGCAGAGGAGCTGCTGCTGGAGGCGCGCTGCTGCGCGG 1618
DB 1561 TGCCGTGCGCGCCAGCAGAGGAGCTGCTGCTGGAGGCGCGCTGCTGCGCGG 1620
QY 1619 ACAACGCGCTGCGCAGCACTACCGGCTGCGAACTGCGAGGCTGCAAGGCTTTTCA 1678
DB 1621 ACAACGCGCTGCGCAGCACTACCGGCTGCGAACTGCGAGGCTGCAAGGCTTTTCA 1680
QY 1679 AGAGAACAGTGCAGAAATATGTTGCTGGCTGGCAAAATAAACTGCCAGTAG 1738
DB 1681 AGAGAACAGTGCAGAAATATGTTGCTGGCTGGCAAAATAAACTGCCAGTAG 1740
QY 1739 ACAAGAGCGTCGAAACCGATGTCAGTACTGCTGATTTCAAGTGTCTCAGTGTGAA 1798
DB 1741 ACAAGAGCGTCGAAACCGATGTCAGTACTGCTGATTTCAAGTGTCTCAGTGTGAA 1800
QY 1799 TGGTAAAGAAAGTTGTCGTAAGTGTGAAAGGAGGAGGCTGCTGCTGCTTCCA 1858
DB 1801 TGGTAAAGAAAGTTGTCGTAAGTGTGAAAGGAGGAGGCTGCTGCTGCTTCCA 1860
QY 1859 AACCAAGAGCGCTTACACAGGAACTTCTCAGCCCTCTCAGCTTCTCCTCAATCT 1918
DB 1861 AACCAAGAGCGCTTACACAGGAACTTCTCAGCCCTCTCAGCTTCTCCTCAATCT 1920
QY 1919 GCATGATGAATGCTTGTGCGAGCTTAAACAGACTCAACCCAGAGATCTTGATTT 1978
DB 1921 GCATGATGAATGCTTGTGCGAGCTTAAACAGACTCAACCCAGAGATCTTGATTT 1980
QY 1979 CCAGATCTGCTCCACGTGACCGAGCTGTCAGGACAGATGCTGAGCATGTGCAACAT 2038
DB 1981 CCAGATCTGCTCCACGTGACCGAGCTGTCAGGACAGATGCTGAGCATGTGCAACAT 2040
QY 2039 TCTACACCTCTGACAGCTTCCATTTGATGATTCAGAAAGTGGGAGGATTCGG 2098
DB 2041 TCTACACCTCTGACAGCTTCCATTTGATGATTCAGAAAGTGGGAGGATTCGG 2100
QY 2099 GATTTACTGATCTCCCAAGAGATCAGACTTACTTATTTGAATCAGCTTTTGGAGC 2158
DB 2101 GATTTACTGATCTCCCAAGAGATCAGACTTACTTATTTGAATCAGCTTTTGGAGC 2160
QY 2159 TGTGTGCTCAGACTTTCCATCAGGTCAACACATGCTGAAAGATAAGTTGTGTTGCA 2218
DB 2161 TGTGTGCTCAGACTTTCCATCAGGTCAACACATGCTGAAAGATAAGTTGTGTTGCA 2220
QY 2219 ATGACTTGTCTGCATCGACTTCACTGCTTCTGATTTGGGAGTGGCTCGACTCTA 2278
DB 2221 ATGACTTGTCTGCATCGACTTCACTGCTTCTGATTTGGGAGTGGCTCGACTCTA 2280
QY 2279 TTAAGACTTTTCTTAAATTTGACAGCTTGAATATCAAGCTTTAGCTTGC 2338
DB 2281 TTAAGACTTTTCTTAAATTTGACAGCTTGAATATCAAGCTTTAGCTTGC 2340
QY 2339 TGTGACACTGAGCATGATCAAGAAAGATGAGGTTAAAGAAACAAAGAGAGTCAAG 2398
DB 2341 TGTGACACTGAGCATGATCAAGAAAGATGAGGTTAAAGAAACAAAGAGAGTCAAG 2400
QY 2399 AGCTATGCAACAGATCAAGAGCTTTAAAGAACCAAGAGTAAAGGAGGCTTGG 2458

Db 2401 AGCTATGCAACAGATCAAGCAGGTTTAAAGACCACCCAGAGTAAGGGACAGGCTCTGG 2460
Qy 2459 AGCCACACGAGTCCAAAGGTCCTGGGTGCCCTGGTAGAACTGAGGAAGATCTGCAACCTCGG 2518
Db 2461 AACCCCAACAGTCCAAAGGTCCTGGTGGCCCTGGTAGAACTGAGGAAGATCTGCAACCTCGG 2520
Qy 2519 GCTCCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTG 2578
Db 2521 GCTCCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTG 2580
Qy 2579 ACAAGCTCTTCTGGACACCCCTACCTCTCTAATCAGGAGCAGTGGAGCAGTGAAGTGCCT 2638
Db 2581 ACAAGCTCTTCTGGACACCCCTACCTCTCTAATCAGGAGCAGTGGAGCAGTGAAGTGCCT 2640
Qy 2639 CCTCTCCTAGCAGCTCTGCTAGCAGCAAAAGGATAGGTTTGGAAAACCTATCATTTCC 2698
Db 2641 CCTCTCCTAGCAGCTCTGCTAGCAGCAAAAGGATAGGTTTGGAAAACCTATCATTTCC 2700
Qy 2699 TGTCTCTCTTAAGAGGAAAGCAGCTCTGTAGAAAGCAAGACTTCTTTTTTTTCG 2758
Db 2701 TGTCTCTCTTAAGAGGAAAGCAGCTCTGTAGAAAGCAAGACTTCTTTTTTTTCG 2760
Qy 2759 GCTCTTTCTCTTAAACCCAGAAAACTTGCAGAGTATTTGTTGGGTTGTGTT 2818
Db 2761 GCTCTTTCTCTTAAACCCAGAAAACTTGCAGAGTATTTGTTGGGTTGTGTT 2820
Qy 2819 TTATATTAGGCAATGGGGATGGGGTGGAGGGGTTATAGTTCATGAGGGTTTCTAA 2878
Db 2821 TTATATTAGGCAATGGGGATGGGGTGGAGGGGTTATAGTTCATGAGGGTTTCTAA 2880
Qy 2879 GAAATTGCTAAACAGCACTTTTGGCAATGCTATCCAGCAGGAGAAAAAGGATAATA 2938
Db 2881 GAAATTGCTAAACAGCACTTTTGGCAATGCTATCCAGCAGGAGAAAAAGGATAATA 2940
Qy 2939 TAACTGTTTTTAAACCTTTCTGGGAAATCCAATTATAGTGTCTTGTATTTAAAAACAA 2998
Db 2941 TAACTGTTTTTAAACCTTTCTGGGAAATCCAATTATAGTGTCTTGTATTTAAAAACAA 3000
Qy 2999 GAACAGCCAAAGGTTGTTCCGAGGATAGATGTCTTAAAGATGGTCCCTTGAAGAT 3058
Db 3001 GAACAGCCAAAGGTTGTTCCGAGGATAGATGTCTTAAAGATGGTCCCTTGAAGAT 3060
Qy 3059 ATGCTTCCTGTATCAAGGTAGTATGTGTGCAACCAAGGAGCAAACTTCTTTTAAAT 3118
Db 3061 ATGCTTCCTGTATCAAGGTAGTATGTGTGCAACCAAGGAGCAAACTTCTTTTAAAT 3120
Qy 3119 TCCTTCTCTCTTATTTTAAACAAATGGTGAAGATGGAGGATTAACCTACAAATCAGACAT 3178
Db 3121 TCCTTCTCTCTTATTTTAAACAAATGGTGAAGATGGAGGATTAACCTACAAATCAGACAT 3180
Qy 3179 GGCAAAACAAATATGCTGTGTTCCATAAACAAGTGCATTTTAAAGTGTGTTCT 3238
Db 3181 GGCAAAACAAATATGCTGTGTTCCATAAACAAGTGCATTTTAAAGTGTGTTCT 3240
Qy 3239 TACTAAGTCTCTTTTAACTCTCTCTTTTATCTATATGGAATATAAGAGGAGGAGTCA 3298
Db 3241 TACTAAGTCTCTTTTAACTCTCTCTTTTATCTATATGGAATATAAGAGGAGGAGTCA 3300
Qy 3299 TGTTAGCAATGACAGCTTAAATATCCCTAGCAGAGCTGTGTTCACTTCCCTGTGATC 3358
Db 3301 TGTTAGCAATGACAGCTTAAATATCCCTAGCAGAGCTGTGTTCACTTCCCTGTGATC 3360
Qy 3359 CCTTCTGAGGTATGGCCCATCCAAGACTTTTAAAGCCATTTCTTGATGGAACAGATCCCTG 3418
Db 3361 CCTTCTGAGGTATGGCCCATCCAAGACTTTTAAAGCCATTTCTTGATGGAACAGATCCCTG 3420
Qy 3419 CCTTGACTCTGAGCTATCCTGAAAGTGGATCAGATTATAAAGTGAATTAAGTAACTG 3478
Db 3421 CCTTGACTCTGAGCTATCCTGAAAGTGGATCAGATTATAAAGTGAATTAAGTAACTG 3480
Qy 3479 TTTTGGTGTGTTCTATCAACCCACAGTTCCTTAACTTCTTCAAGTTATAGTAAC 3538
Db 3481 TTTTGGTGTGTTCTATCAACCCACAGTTCCTTAACTTCTTCAAGTTATAGTAAC 3540

Qy 3539 TGACTGGTATATATTCATTGAGAGCGCCATAAAGTCAAGTTCAGTATTTGATCCCTAGATAAG 3598
Db 3541 TGACTGGTATATATTCATTGAGAGCGCCATAAAGTCAAGTTCAGTATTTGATCCCTAGATAAG 3600
Qy 3599 AACATGCAAAATCAGCAGGAACCTGGTCAACAGGTAAGCACCAGGCAACAATAAGGATTTT 3658
Db 3601 AACATGCAAAATCAGCAGGAACCTGGTCAACAGGTAAGCACCAGGCAACAATAAGGATTTT 3660
Qy 3659 TATAGATATAAATTTAATTTTGTTA--TTGGTTAAGGAGA-CAATTTTGGAGAGCAAGCA 3715
Db 3661 TATAGATATAAATTTAATTTTGTAAATTTGGTAAATTTAGGAGACCAATTTTGGAGAGCAAGCA 3720
Qy 3715 AA---TCITTTTAAAAAATAGTATGAATGTGAATCTAGAAAAAGATTTTAAAAAATAGTAT 3772
Db 3721 AATCTTCTTTTAAAAAATAGTATGAATGTGAATCTAGAAAAAGATTTTAAAGAAATAGTAT 3780
Qy 3773 GAGTGTGAGTACTAGGAAGGAT 3794
Db 3781 GAGTGTGAGTACTAGGAAGGAT 3802

RESULT 4

US-10-414-080-1

; Sequence 1, Application US/10414080

; Publication No. US2003020288A1

; GENERAL INFORMATION:

; APPLICANT: MULLICAN, SHANNON E.

; APPLICANT: CONNELLY, ORLA M.

; APPLICANT: MILBRANDT, JEFFREY

; TITLE OF INVENTION: NOR-1 AND NUR77 NUCLEAR RECEPTORS AS TARGETS FOR

; FILE REFERENCE: P02454US1

; CURRENT APPLICATION NUMBER: US/10/414,080

; PRIOR FILING DATE: 2003-04-15

; PRIOR APPLICATION NUMBER: 60/373,238

; PRIOR FILING DATE: 2002-04-17

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 3802

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-414-080-1

Query Match 98.1%; Score 3722; DB 17; Length 3802;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

Qy 1 ATAAATGACGTCCGAGAGAGCGAGAAACGGCGAGCGGGAGAGCGGAGTCTCTGCCT 60
Db 1 ATAAATGACGTCCGAGAGAGCGAGAAACGGCGAGCGGGAGAGCGGAGTCTCTGCCT 60
Qy 61 CCGCGCCCCACCCCTCCAGCTCCTCTCCTCCGCTCCCATACACAGACGGCTCA 120
Db 61 CCGCGCCCCACCCCTCCAGCTCCTCTCCTCCGCTCCCATACACAGACGGCTCA 120
Qy 121 CACCGCTCCCTCACTCGCACACAGACACAAGCGCGCACACAGGCTCCGACACAC 178
Db 121 CACCGCTCCCTCACTCGAACACACAGACACAAGCGCGCACACAGGCTCCGACACAC 180
Qy 179 ACTTCGCTCTCCCGCGGCTCAACCCCTCTTGGCCCTGAGCCCTTGGCCGGTGGCGG 238
Db 181 ACTTCGCTCTCCCGCGGCTCAACCCCTCTTGGCCCTGAGCCCTTGGCCGGTGGCGG 240
Qy 239 GCGCGCAGCTGAGCCCTCCCGGGCTCACTTCGACGGCTGACGGTCCCGCAGTGGC 298
Db 241 GCGCGCAGCTGAGCCCTCCCGGGCTCACTTCGACGGCTGACGGTCCCGCAGTGGC 300
Qy 299 CGTGAGGTGGGAAACAGCGGGGGCATCTCTCCCTTGTGTACAGCCCAAGCCAGGACGC 358
Db 301 CGTGAGGTGGGAAACAGCGGGGGCATCTCTCCCTTGTGTACAGCCCAAGCCAGGACGC 360

Db 2521 GCCTCAGCGCATCTTCTACCTGAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTG 2580
QY 2579 ACAAGCTCTTCTGGACACCCCTACCTTTCTAATTCAGAGCAGTGAGCAGTGCAGCTGCT 2638
Db 2581 ACAAGCTCTTCTGGACACCCCTACCTTTCTAATTCAGAGCAGTGAGCAGTGCAGCTGCT 2640
QY 2639 CCTCTCCCTAGCACCTGCTTGTCTACGACGAAAGGATAGTTTGGAAACCTATCATTTCC 2698
Db 2641 CCTCTCCCTAGCACCTGCTTGTCTACGACGAAAGGATAGTTTGGAAACCTATCATTTCC 2700
QY 2699 TGTCTCTTCCCTTAAGAGAAAAGCAGCTCCTGTAGAAAAGCAAAAGACTTCTTTTTCGTG 2758
Db 2701 TGTCTCTTCCCTTAAGAGAAAAGCAGCTCCTGTAGAAAAGCAAAAGACTTCTTTTTCGTG 2760
QY 2759 GCTCTTTTCCCTTAACACTAAAGCCAGAAAACCTTCAGAGATATGTGTGGGGTGTGTGT 2818
Db 2761 GCTCTTTTCCCTTAACACTAAAGCCAGAAAACCTTCAGAGATATGTGTGGGGTGTGTGT 2820
QY 2819 TTATATTTAGGCATTTGGGGATGGGGTGGAGGGGTATAGTTTCATGAGGGTTCCTAA 2878
Db 2821 TTATATTTAGGCATTTGGGGATGGGGTGGAGGGGTATAGTTTCATGAGGGTTCCTAA 2880
QY 2879 GAAATTTGCTTAACAAAGCACTTTTGGCAATGCTATCCACAGCAGGAAAAAGGATAATA 2938
Db 2881 GAAATTTGCTTAACAAAGCACTTTTGGCAATGCTATCCACAGCAGGAAAAAGGATAATA 2940
QY 2939 TAACTGTTTTAAACACTTTTCTGGGGAATCCAAATATAGTTGCTTTGTAATTTAAACAA 2998
Db 2941 TAACTGTTTTAAACACTTTTCTGGGGAATCCAAATATAGTTGCTTTGTAATTTAAACAA 3000
QY 2999 GAACAGCCAAAGGTTGTCGCGCAGGATAGTGTCTTAAAGATTGGTCCCTTCAAAAT 3058
Db 3001 GAACAGCCAAAGGTTGTCGCGCAGGATAGTGTCTTAAAGATTGGTCCCTTCAAAAT 3060
QY 3059 ATGCTTCTCTGTATCAAAAGTAGTATGTGTGTCMAAAGCAGGCAAACTTCTTTTAAAT 3118
Db 3061 ATGCTTCTCTGTATCAAAAGTAGTATGTGTGTCMAAAGCAGGCAAACTTCTTTTAAAT 3120
QY 3119 TCCTTCTTCTTTATTTTAAACAAATGGTGAAGATGAGGATTAAGTCAAAATCAGACAT 3178
Db 3121 TCCTTCTTCTTTATTTTAAACAAATGGTGAAGATGAGGATTAAGTCAAAATCAGACAT 3180
QY 3179 GCGAAACAAATAGCTGTTGCTTCCATAACAGTGCATTTTAAAGTGTGCT 3238
Db 3181 GCGAAACAAATAGCTGTTGCTTCCATAACAGTGCATTTTAAAGTGTGCT 3240
QY 3239 TACTAAGTCTGTTTATTAACCTCTCCTTTATTTCTATATGAAATAAAGAGGAGTGCA 3298
Db 3241 TACTAAGTCTGTTTATTAACCTCTCCTTTATTTCTATATGAAATAAAGAGGAGTGCA 3300
QY 3299 TGTAGCAATGACACGTTTAAATATCCCTAGCAGAGCTGTGTACCTTCCCTGTGATC 3358
Db 3301 TGTAGCAATGACACGTTTAAATATCCCTAGCAGAGCTGTGTACCTTCCCTGTGATC 3360
QY 3359 CCTTCTGAGGTATGCGCCATCCAAGACTTTTAGGCCATTTCTGTAGAACCAAGATCCCTG 3418
Db 3361 CCTTCTGAGGTATGCGCCATCCAAGACTTTTAGGCCATTTCTGTAGAACCAAGATCCCTG 3420
QY 3419 CCTGACTCTCCAGCTATCTCTGAAAGTGATCAGATTATATAAACTGGATTACATGTAAC 3478
Db 3421 CCTGACTCTCCAGCTATCTCTGAAAGTGATCAGATTATATAAACTGGATTACATGTAAC 3480
QY 3479 TTTTGGTGTGTCTATCAACCCCAACAGGTTCCCTAAACTGTGCTTCAGTTATAGTAAC 3538
Db 3481 TTTTGGTGTGTCTATCAACCCCAACAGGTTCCCTAAACTGTGCTTCAGTTATAGTAAC 3540
QY 3539 TGACTGGTATATTCATTGAGAGCGCCATAAGTCAGTTGAGTATTTGATCCCTAGATAAG 3598
Db 3541 TGACTGGTATATTCATTGAGAGCGCCATAAGTCAGTTGAGTATTTGATCCCTAGATAAG 3600
QY 3599 AACATGCAATCAGCAGGAACCTGGTCTATACAGGGTAAGCACCAGGGCAACAATAAGGATTT 3658
Db 3601 AACATGCAATCAGCAGGAACCTGGTCTATACAGGGTAAGCACCAGGGCAACAATAAGGATTT 3660

QY 3659 TATAGATATAATTTAATTTTGTGTTA--TTGGTTAAGGAGA--CAATTTTGGAGAGCAAGCA 3715
Db 3661 TATAGATATAATTTAATTTTGTGTTAATTTGGTAAAGGAGACCAATTTTGGAGAGCAAGCA 3720
QY 3716 AA---TCTTTTAAATAATAGTATGAATCTAGAAAGATTTTAAAAAATAGTAT 3772
Db 3721 AATCTCTTTTAAATAATAGTATGAATCTAGAAAGATTTTAAAGAAATAGTAT 3780
QY 3773 GAGTGTGAGTACTAGAAAGGAT 3794
Db 3781 GAGTGTGAGTACTAGAAAGGAT 3802

RESULT 5

US-10-659-004-115
; Sequence 115, Application US/10659004
; Publication No. US20050048507A1
; GENERAL INFORMATION:
; APPLICANT: Zhong et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-608
; CURRENT APPLICATION NUMBER: US/10/659,004
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 115
; LENGTH: 3802
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (732)..(2609)
US-10-659-004-115

Query Match 98.1%; Score 3722; DB 19; Length 3802;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY 1 ATAAATGACGTCCGAGAGAGCGGAGCAACGCGCAGCGGAGCGGAGTCTCTGCGCT 60
Db 1 ATAAATGACGTCCGAGAGAGCGGAGCAACGCGCAGCGGAGCGGAGTCTCTGCGCT 60
QY 61 CCGCGCCCCCACCCTCCAGCTCCCTGCTCTCTCCGCTCCCATACACAGACGCGCTCA 120
Db 61 CCGCGCCCCCACCCTCCAGCTCCCTGCTCTCTCTCTCTCCCATACACAGACGCGCTCA 120
QY 121 CACCGCGCTCCCTCACTCCGCGCACACAGACACAAGCGCGCACACAGGCTCCG--CACACAC 178
Db 121 CACCGCGCTCCCTCACTCCGCGCACACAGACACAAGCGCGCACACAGGCTCCGCGCACACAC 180
QY 179 ACTTGTCTCTCCGCGCGCTCAACCCCTCTTGTGCTGAGCCCTTGTCCGCGTGCAGCGCGG 238

181	ACTTGGCTCTCCGGCGCTCACACCTCTTGGCCCTGAGCCCTTCCGGTGGACCGG	240	1261	TGAAGCGGTGCCCCACGCTGGCCGGCGCGCTTCCCGCTTTCACATTTCAAGCCCTCGC	1320
239	CGCCGAGCTGAGACCGCCCTCCCGGCTCACTTTGCAACGCTGACGCTGCGGAGTGC	298	1319	CGCGCATCCCG	1378
241	CGCCGAGCTGAGACCGCCCTCCCGGCTCACTTTGCAACGCTGACGCTGCGGAGTGC	300	1321	CGCGCATCCCG	1380
299	CGTGGAGTGGAAACAGCGGGGCGCATCTCCCGCTTGGTCAAGCCCAAGCCAGGACGC	358	1379	CGCTCGCGCGCTCAGCCTTCCCGCTTGGGAGCGCGCGCGCGCGCGCGCGCGCGCG	1438
301	CGTGGAGTGGAAACAGCGGGGCGCATCTCCCGCTTGGTCAAGCCCAAGCCAGGACGC	360	1381	CGCTCGCGCGCTCAGCCTTCCCGCTTGGGAGCGCGCGCGCGCGCGCGCGCGCGCG	1440
359	CGCGAACCTCTCGCTGCTCTCCATGAGTGGGATCGAGCATCCCGCCAGCGCG	418	1439	CGTTGAGAGCACCCTGACGCTCGCGCTGGCAGAGGCGCGCGCGCGCGCGCGCTTC	1498
361	CGCGAACCTCTCGCTGCTCTCCATGAGTGGGATCGAGCATCCCGCCAGCGCG	420	1441	CGTTGAGAGCACCCTGACGCTCGCGCTGGCAGAGGCGCGCGCGCGCGCGCTTC	1500
419	CTCACCGCTCCGGGAGCGCTGGGCTTGTACACCGCAGCCCTTCCGGGACAGAGTGT	478	1499	CGCTCTCGSCCTCAGCCCTCCCTACCGCGTTCAGCCCTGCTGGGCGAGAGTCCAGCC	1558
421	CTCACCGCTCCGGGAGCGCTGGGCTTGTACACCGCAGCCCTTCCGGGACAGAGTGT	480	1501	CGCTCTCGSCCTCAGCCCTCCCTACCGCGTTCAGCCCTGCTGGGCGAGAGTCCAGCC	1560
479	GACTCCCCCAGTGAGATTTTCGGGACAGCTCTCTAGAACTGCTCTAAGAGCGAAC	538	1559	TGCCGTCCCGCGCAGCAGGAGCTCGTGTCTGGCGAGGCGAGTGTGCGGTGGGGG	1618
481	GACTCCCCCAGTGAGATTTTCGGGACAGCTCTCTAGAACTGCTCTAAGAGCGAAC	540	1561	TGCCGTCCCGCGCAGCAGGAGCTCGTGTCTGGCGAGGCGAGTGTGCGGTGGGGG	1620
539	CGCCACAGCACTCAAGCCCACTGCGGAAGAGGCGAGCCCGGCAAGCCCGGCGCTGAGC	598	1619	ACACGCGCGCTGCCAGCACTACGGGTGCGAACCTCGAGGCGCTGCAAGGGCTTTTCA	1678
541	CGCCACAGCACTCAAGCCCACTGCGGAAGAGGCGAGCCCGGCAAGCCCGGCGCTGAGC	600	1621	ACACGCGCGCTGCCAGCACTACGGGTGCGAACCTCGAGGCGCTGCAAGGGCTTTTCA	1680
599	CTGGACCTTACGCTGCGGGGAGCACTGCGGGGCTTGGCTCGCGGAGCTCGGCTC	658	1679	AGAGAACGTCGAGAAATATGCTTTCCTGGCAATAAATAATGCTTTCCTGCTAGTAG	1738
601	CTGGACCTTACGCTGCGGGGAGCACTGCGGGGCTTGGCTCGCGGAGCTCGGCTC	660	1681	AGAGAACGTCGAGAAATATGCTTTCCTGGCAATAAATAATGCTTTCCTGCTAGTAG	1740
659	CTCCTACCTCTCAGCTCGCTGGAGAGACCCCGAGCCCACTTTCAGCGCGCAAGAT	718	1739	ACAGAGAGCTCGAAACCGATGTCAGTGTGCTGATTTTCAGAGTGTCTCAGTGTGGAA	1798
661	CTCCTACCTCTCAGCTCGCTGGAGAGACCCCGAGCCCACTTTCAGCGCGCAAGAT	720	1741	ACAGAGAGCTCGAAACCGATGTCAGTGTGCTGATTTTCAGAGTGTCTCAGTGTGGAA	1800
719	ACCTCTCAGATATGCTGCTGCTCAAGCCCAATATAGCCCTTCCCTCCAGTTCAGTGT	778	1799	TGTTAAAGAGTGTGCTGACAGATAGTCTGAAAGGAGGAGAGAGTGTCTGCTTCCA	1858
721	ACCTCTCAGATATGCTGCTGCTCAAGCCCAATATAGCCCTTCCCTCCAGTTCAGTGT	780	1801	TGTTAAAGAGTGTGCTGACAGATAGTCTGAAAGGAGGAGAGAGTGTCTGCTTCCA	1860
779	ATGGGGGAGACATACAGCTCGGATACACAGGAGATCATGAACCCGACTACACCA	838	1859	AACCAAGAGCGCTTCAACAGGAGCTTCTCAGCCCTCTCCACCTTCTCTCCAATCT	1918
781	ATGGGGGAGACATACAGCTCGGATACACAGGAGATCATGAACCCGACTACACCA	840	1861	AACCAAGAGCGCTTCAACAGGAGCTTCTCAGCCCTCTCCACCTTCTCTCCAATCT	1920
839	AGCTGACATGAGCTTGGCAGCACTGAGATACGGCTACAGCACCACTGCTGCCCA	898	1919	GCATGATGATGCTTGTGCTGAGCTTTACAGACTCAACACCCAGAGATCTTGAATATT	1978
841	AGCTGACATGAGCTTGGCAGCACTGAGATACGGCTACAGCACCACTGCTGCCCA	900	1921	GCATGATGATGCTTGTGCTGAGCTTTACAGACTCAACACCCAGAGATCTTGAATATT	1980
899	GCATGATGCTTGGGAGGCTACTCGAGCACTACGAACTCAAGCCCTTCTGCGTGT	958	1979	CCAGATCTGCTCCACTGACCGAGCTGTCAGGCGACAGATGCTGAGCATGTGCAACAT	2038
901	GCATGATGCTTGGGAGGCTACTCGAGCACTACGAACTCAAGCCCTTCTGCGTGT	960	1981	CCAGATCTGCTCCACTGACCGAGCTGTCAGGCGACAGATGCTGAGCATGTGCAACAT	2040
959	ACCAATGACAGCGCTTGCATCAAGTGGAGGAGGCGCGCGCGCGCGCGCGCGCGCG	1018	2039	TCTACAACTCTGACAGCGCTTCCATTTGATGATCCAGAGCTGGGCGAGAAAGATTCGGG	2098
961	ACCAATGACAGCGCTTGCATCAAGTGGAGGAGGCGCGCGCGCGCGCGCGCGCGCG	1020	2041	TCTACAACTCTGACAGCGCTTCCATTTGATGATCCAGAGCTGGGCGAGAAAGATTCGGG	2100
1019	ATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	1078	2099	GATTTACTGATCTCCCAAGAGATCAGACATTTATTTGAATTCAGCTTTTGGAGC	2158
1021	ATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	1080	2101	GATTTACTGATCTCCCAAGAGATCAGACATTTATTTGAATTCAGCTTTTGGAGC	2160
1079	TTCCTCAGCTCCAGCCCGGAGGAGGCTGCTGCGGAGCTGCTGCTGCTGCTGCTGCT	1138	2159	TGTTGCTCTCAGATCTTCCATCAGGTCAAACCTGCTGAAGATAAGTTTGTGTTGCA	2218
1081	TTCCTCAGCTCCAGCCCGGAGGAGGCTGCTGCGGAGCTGCTGCTGCTGCTGCTGCT	1140	2161	TGTTGCTCTCAGATCTTCCATCAGGTCAAACCTGCTGAAGATAAGTTTGTGTTGCA	2220
1139	AGTCCCAACCTCCACCCAGCAGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCG	1198	2219	ATGGAGTGTCTGCTGATGCTTCCAGTGTCTGCTGAGATTTGGGAGTGGCTCGACTCTA	2278
1141	AGTCCCAACCTCCACCCAGCAGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCG	1200	2221	ATGGAGTGTCTGCTGATGCTTCCAGTGTCTGCTGAGATTTGGGAGTGGCTCGACTCTA	2280
1199	ACGAGGACCTGCTCGCGCGCGCGCTGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCT	1258	2279	TTAAAGACTTTTCTTAAATTTGAGAGCTGAACTTATATCAAGCTTTAGCTTCC	2338
1201	ACGAGGACCTGCTCGCGCGCGCGCTGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCT	1260	2281	TTAAAGACTTTTCTTAAATTTGAGAGCTGAACTTATATCAAGCTTTAGCTTCC	2340
1259	TGAAGCGGTGCCCCACGCTGGCCGGCGCGCTTCCCGCTTTCACATTTCAAGCCCTCGC	1318	2339	TGTCAGCTGAGAGTATCAAGAAAGACATGCGGTAAAGAAACCAAGAGAGTGAAG	2398
			2341	TGTCAGCTGAGAGTATCAAGAAAGACATGCGGTAAAGAAACCAAGAGAGTGAAG	2400

Qy	2399	AGCTATGCAACAAGATCACAAAGCAGTTTTAAAGACCAACAGATTAAGGGACAGGCTCTGG	2458
Db	2401	AGCTATGCAACAAGATCACAAAGCAGTTTTAAAGACCAACAGATTAAGGGACAGGCTCTGG	2460
Qy	2459	AGCCCAACGAGTCCAAAGTCTCGGTGCCCTGGTGAACCTGAGGAAGATCTGCACCCCTGG	2518
Db	2461	AACCCAAACGAGTCCAAGTCTCGGTGCCCTGGTGAACCTGAGGAAGATCTGCACCCCTGG	2520
Qy	2519	GCCTCCAGCGCATCTTCTACCTGAAGCTGGAAGCATTTGGTGTCTCCACCTTCCATCATTTG	2578
Db	2521	GCCTCCAGCGCATCTTCTACCTGAAGCTGGAAGCATTTGGTGTCTCCACCTTCCATCATTTG	2580
Qy	2579	ACAAGCTCTTCTGGACACCTTACCTTTCTAATCAGGACAGTGGAGCAGTAGCTGCCT	2638
Db	2581	ACAAGCTCTTCTGGACACCTTACCTTTCTAATCAGGACAGTGGAGCAGTAGCTGCCT	2640
Qy	2639	CCCTCTCTAGCACCTGCTTCTACGACGACAAAGGATAGGTTTTGGAAACCTATCATTTCC	2698
Db	2641	CCCTCTCTAGCACCTGCTTCTACGACGACAAAGGATAGGTTTTGGAAACCTATCATTTCC	2700
Qy	2699	TGTCCTTCTTAAAGAGAAAGCAGCTCCTGTAGAAGCAAGACCTTCTTTTTTTCTG	2758
Db	2701	TGTCCTTCTTAAAGAGAAAGCAGCTCCTGTAGAAGCAAGACCTTCTTTTTTTCTG	2760
Qy	2759	GCTCTTTTCTTACAACCTTAAAGCCAGAAACCTTGCAGAGTATTGTGTTGGGGTTGTGTT	2818
Db	2761	GCTCTTTTCTTACAACCTTAAAGCCAGAAACCTTGCAGAGTATTGTGTTGGGGTTGTGTT	2820
Qy	2819	TTATATTTAGGCATTTGGGGATGCGGTGGAGGGGGTTATAGTTCATGAGGGTTTCTAA	2878
Db	2821	TTATATTTAGGCATTTGGGGATGCGGTGGAGGGGGTTATAGTTCATGAGGGTTTCTAA	2880
Qy	2879	GAATTTGCTTAACAAAGCACCTTTTGGACAAATGCTATCCCAGCAGGAAAAAAGGATAATA	2938
Db	2881	GAATTTGCTTAACAAAGCACCTTTTGGACAAATGCTATCCCAGCAGGAAAAAAGGATAATA	2940
Qy	2939	TAACCTGTTTTAAACCTCTTCTGGGAAATCCAATTATAGTTGCTTGTATTTAAAAACAA	2998
Db	2941	TAACCTGTTTTAAACCTCTTCTGGGAAATCCAATTATAGTTGCTTGTATTTAAAAACAA	3000
Qy	2999	GAACAGCAAGGGTGTGTCGCAGGGTAGGATGTGCTTAAAGATTGGTCCCTTGAAAT	3058
Db	3001	GAACAGCAAGGGTGTGTCGCAGGGTAGGATGTGCTTAAAGATTGGTCCCTTGAAAT	3060
Qy	3059	ATGCTTCTGTATCAAAAGTACGTATGTGGTGCAAAACAGGCAGAACTTCTTTTAAAT	3118
Db	3061	ATGCTTCTGTATCAAAAGTACGTATGTGGTGCAAAACAGGCAGAACTTCTTTTAAAT	3120
Qy	3119	TCCTTCTTCTTTATTTTAAACAATGGTGAAGATGGAGGATTAACCTACAAATCAGACAT	3178
Db	3121	TCCTTCTTCTTTATTTTAAACAATGGTGAAGATGGAGGATTAACCTACAAATCAGACAT	3180
Qy	3179	GGCAAAACAATAATGGCTGTGCTTCCATAAACAAGTGCAAATTTTAAAGTGCTGTCT	3238
Db	3181	GGCAAAACAATAATGGCTGTGCTTCCATAAACAAGTGCAAATTTTAAAGTGCTGTCT	3240
Qy	3239	TACTAAGTCTGTGTTATTAACCTCTCTTTTATCTATATGAAATAAAAAAGGAGGACGTCA	3298
Db	3241	TACTAAGTCTGTGTTATTAACCTCTCTTTTATCTATATGAAATAAAAAAGGAGGACGTCA	3300
Qy	3299	TGTTAGCAANTGACAGCTTAATATCCCTACGACAGGCTGTGTTCACTTCCCTGTCGATC	3358
Db	3301	TGTTAGCAANTGACAGCTTAATATCCCTACGACAGGCTGTGTTCACTTCCCTGTCGATC	3360
Qy	3359	CCCTTCTGAGGTATGGCCCATCCAAGACTTTTAGGCCAATCTTGATGGAAACAGATCCCTG	3418
Db	3361	CCCTTCTGAGGTATGGCCCATCCAAGACTTTTAGGCCAATCTTGATGGAAACAGATCCCTG	3420
Qy	3419	CCCTGACTGTCAGATATCTGAAAGTGGATACAGATTATAAATCGGATTAATATGTAACCTG	3478
Db	3421	CCCTGACTGTCAGATATCTGAAAGTGGATACAGATTATAAATCGGATTAATATGTAACCTG	3480

Qy	3479	TTTTGGTGTGGTCTATCAACCCCAACAGAGTTCCTAAACTTGCCTCAGTTATAGTAAC	3538
Db	3481	TTTTGGTGTGGTCTATCAACCCCAACAGAGTTCCTAAACTTGCCTCAGTTATAGTAAC	3540
Qy	3539	TGACTGTGATATTTCAATCAGAAGCGCCATAAGTCAGTTCAGTATTTTGATCCCTAGATAAG	3598
Db	3541	TGACTGTGATATTTCAATCAGAAGCGCCATAAGTCAGTTCAGTATTTTGATCCCTAGATAAG	3600
Qy	3599	AACATGCCAATCAGCAGGACCTGGTCATACAGGGTAAGCACCGAGGACCAATAGGATTTT	3658
Db	3601	AACATGCCAATCAGCAGGACCTGGTCATACAGGGTAAGCACCGAGGACCAATAGGATTTT	3660
Qy	3659	TATAGATATATTTTAAATTTTTTGTTA--TTGGTTAAGGAGA-CAATTTTGGAGAGCAAGCA	3715
Db	3661	TATAGATATATTTTAAATTTTTTGTTA--TTGGTTAAGGAGACCAATTTTGGAGAGCAAGCA	3720
Qy	3716	AA---TCCTTTTAAAAAATAGTATGAATCTCGAATCTCGAAGAGATTTAAAAAATAGTAT	3772
Db	3721	AATCTCTTTTTAAAAAATAGTATGAATCTCGAATCTCGAAGAGATTTAAAAAATAGTAT	3780
Qy	3773	GAGTGTGAGTACTAGGAAGGAT	3794
Db	3781	GAGTGTGAGTACTAGGAAGGAT	3802

RESULT 6

```

US-10-247-671-105
: Sequence 105, Application US/10247671
: Publication No. US20030194721A1
: GENERAL INFORMATION:
: APPLICANT: Mikita, Thomas
: APPLICANT: Shifman, Dov
: APPLICANT: Porter, Gordon, J.
: APPLICANT: Kaser, Matthew R.
: TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
: FILE REFERENCE: PA-0050 US
: CURRENT APPLICATION NUMBER: US/10/247,671
: CURRENT FILING DATE: 2002-09-18
: PRIOR APPLICATION NUMBER: 60/323,784
: PRIOR FILING DATE: 2001-09-19
: NUMBER OF SEQ ID NOS: 186
: SOFTWARE: PERL Program
: SEQ ID NO 105
: LENGTH: 5828
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No. US20030194721A1 349343.3
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 1012-1074, 4044-4074, 5807
: OTHER INFORMATION: a, t, c, g, or other
US-10-247-671-105

```

	Query Match	97.0%	Score 3679.2	DB 16	Length 5828	
	Best Local Similarity	98.1%	Pred. No. 0			
	Matches 3728	Conservative	0	Mismatches 66	Indels 7	Gaps 4
Qy	1	ATAAATGACGTGCCGAGAGACGAGCGAAGCGCGAGCGGGAGAGCGGAGTCTCTGCGCT	60			
Db	1	ATAAATGACGTGCCGAGAGACGAGCGAAGCGCGAGCGGGAGAGCGGAGTCTCTGCGCT	60			
Qy	61	CCGCGCCCCACCCCTCCAGCTCTGCTCTCTCGCTCCCATACACAGACGCGCTCA	120			
Db	61	CCGCGCCCCACCCCTCCAGCTCTGCTCTCTCGCTCCCATACACAGACGCGCTCA	120			
Qy	121	CACCGGTCCTCTCATCTGCAACACACAGACAACAGCGCGGCACACAGGCTCG--CACACAC	178			
Db	121	CACCGGTCCTCTCATCTGAAACACACAGACAACAGCGCGCACACAGGCTCCGCACACACAC	180			
Qy	179	ACTTTCGTCTTCCCGCGGCTTCACACCCCTCTTGCCCTTGAGCCCTTTCGCGTGAGGCGCG	238			

Db 181 ACTTCGCTCTCCGCGCGCTACACCCCTCTTGGCCCTGAGCCCTTGCCTGCGGTGCGAGCGG 240
 Qy 239 CGCCGAGCTGAGCGCCCTCCGCGCTCACTTTGCAACGCTGACGCTGCGCGAGTGGC 298
 Db 241 CGCGGAGCTGAGCGCCCTCCGCGCTCACTTTGCAACGCTGAGCTGCGCGAGTGGC 300
 Qy 299 CGTGAAGTGGAAACGAGCGCGGCTATCTCCCTCTGCTCAAGCCCAAGCCAGCGAGCGCC 358
 Db 301 CGTGAAGTGGAAACGAGCGCGGCTATCTCCCTCTGCTCAAGCCCAAGCCAGCGAGCGCC 360
 Qy 359 CGCGGAACCTCTCGGCTGCTCTCCATGAGTCCGGATCGCAGCATCCCCACAGCGC 418
 Db 361 CGCGGAACCTCTCGGCTGCTCTCCATGAGTCCGGATCGCAGCATCCCCACAGCGC 420
 Qy 419 CTCACGCTCCGCGGAGCGCTGGCTTGTACAACGCGAGCCCTTCCGGGACGAGCTGT 478
 Db 421 CTCACGCTCCGCGGAGCGCTGGCTTGTACAACGCGAGCCCTTCCGGGACGAGCTGT 480
 Qy 479 GACTCCCGCCAGTGCAGATTTGGGACAGCTCTCTAGAACTCTGCTCTTAAAGACGGAAC 538
 Db 481 GACTCCCGCCAGTGCAGATTTGGGACAGCTCTCTAGAACTCTGCTCTTAAAGACGGAAC 540
 Qy 539 CGCCACAGCACTCAAAAGCCCACTGCGGAAGAGGCGAGCCCGGCAAGCCCGGCGCTGAGC 598
 Db 541 CGCCACAGCACTCAAAAGCCCACTGCGGAAGAGGCGAGCCCGGCAAGCCCGGCGCTGAGC 600
 Qy 599 CTGGACCTTAGCGGTGCGGCGAGCACTGCGGCGCTTTCGCTCGCGGAGCTCGGCTC 658
 Db 601 CTGGACCTTAGCGGTGCGGCGAGCACTGCGGCGCTTTCGCTCGCGGAGCTCGGCTC 660
 Qy 659 CTCTCACTCTCAGCCTTCGCTGAGAGACCCCGGCAAGCCCACTTCAAGCGCGCAAGAT 718
 Db 661 CTCTCACTCTCAGCCTTCGCTGAGAGACCCCGGCAAGCCCACTTCAAGCGCGCAAGAT 720
 Qy 719 ACCCTCAGATATGCGCTGCGTCCAAAGCCCAATATAGCCCTTCCCTTCCAGTTCCAGTT 778
 Db 721 ACCCTCAGATATGCGCTGCGTCCAAAGCCCAATATAGCCCTTCCCTTCCAGTTCCAGTT 780
 Qy 779 ATGGCGCGCAGATACAGCTCGGATACACACGAGATCATGAACCCCGACTACACCA 838
 Db 781 ATGGCGCGCAGATACAGCTCGGATACACACGAGATCATGAACCCCGACTACACCA 840
 Qy 839 AGCTGACATGGAACCTTTGGCAGCACTGAGATACGGCTACAGCAACGCTCCCTGCCCA 898
 Db 841 AGCTGACATGGAACCTTTGGCAGCACTGAGATACGGCTACAGCAACGCTCCCTGCCCA 900
 Qy 899 GCATCAGTACTTGGTGGAGGCTACTCGAGCACTACGAACCTCAAGCCCTTCTGGTGT 958
 Db 901 GCATCAGTACTTGGTGGAGGCTACTCGAGCACTACGAACCTCAAGCCCTTCTGGTGT 960
 Qy 959 ACCAAATGACGCGCCCTTGATCAAAAGTGGAGGAGGGCGGCGCCAGCTACCATOACC 1018
 Db 961 ACCAAATGACGCGCCCTTGATCAAAAGTGGAGGAGGGCGGCGCCAGCTNNNNNNNN 1020
 Qy 1019 ATCACCACCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1078
 Db 1021 NNN 1080
 Qy 1079 TTCTTCCAGCTTCCAGCCCGGAGGAGGAGTGTGCGCCAGCACCTCCATGTACTTCAAGC 1138
 Db 1081 TTCTTCCAGCTTCCAGCCCGGAGGAGGAGTGTGCGCCAGCACCTCCATGTACTTCAAGC 1140
 Qy 1139 AGTCCCAACCGTCCACCCCAACCAACCGCGCTTCCCGCCGAGCGGGGGCGGTATGGG 1198
 Db 1141 AGTCCCAACCGTCCACCCCAACCAACCGCGCTTCCCGCCGAGCGGGGGCGGTATGGG 1200
 Qy 1199 ACGAGGCACTGCGCTCGGCGCCCGGCTGATCGCACCGGCGCTTCTGGACCGCGCGA 1258
 Db 1201 ACGAGGCACTGCGCTCGGCGCCCGGCTGATCGCACCGGCGCTTCTGGACCGCGCGA 1260
 Qy 1259 TGAAGGCGGTCCCCACGGTGGCGGCGGCTTCCCGCTTCCCTTCAAGCCCTCGC 1318
 Db 1261 TGAAGGCGGTCCCCACGGTGGCGGCGGCTTCCCGCTTCTTCCACTTCAAGCCCTCGC 1320

Qy 1319 CGCGCATCCCGCGCGCCAGCCCGCGCGGCGGCCACACCTCGCTCGCTACGACCGAGCG 1378
 Db 1321 CGCGCATCCCGCGCGCCAGCCCGCGCGGCGGCCACACCTCGCTCGCTACGACCGAGCG 1380
 Qy 1379 CGCGTCCGCGCTCAGCTGCGCTGGAGCGCGAGCGCGCGCGGCGGCGAGCGCGCG 1438
 Db 1381 CGCGTCCGCGCTCAGCTGCGCTGGAGCGCGAGCGCGCGCGGCGGCGAGCGCGCG 1440
 Qy 1439 CGTTGAGAGCAACCGTACGCGCTGCGTGGCCAAAGAGGCGGCGCGCGCTGGCTTCC 1498
 Db 1441 CGTTGAGAGCAACCGTACGCGCTGCGTGGCCAAAGAGGCGGCGCGCGCTGGCTTCC 1500
 Qy 1499 CGCTCTCGGCTCAGCGCTTCCCTACCGCTCGCTGGGCGAGAGTCCAGCG 1558
 Db 1501 CGCTCTCGGCTCAGCGCTTCCCTACCGCTCGCTGGGCGAGAGTCCAGCG 1560
 Qy 1559 TGGCGTCCGCGCCAGCAGGAG-CTGGTCTGCTGGCGAGGCGACGTGTGCGGTGGGG 1617
 Db 1561 TGGCGTCCGCGCCAGCAGGAGTCTGCTGCTGGCGAGGCGACGTGTGCGGTGGGG 1620
 Qy 1618 GACAAACCGCG-CTGCGCAGCACTACGCGCTGCGAAACCTGCGAGGCTGCAAGGGCTTTT 1676
 Db 1621 GACAAACCGCGCACTGCGCAGCACTACGCGCTGCGAAACCTGCGAGGCTGCAAGGGCTTTT 1680
 Qy 1677 CAAGAGAACAGTGCAGAAAAATGCAAAATATGTTGCTGGCAAAATGCAAACTGCCAGT 1736
 Db 1681 CAAGAGAACAGTGCAGAAAAATGCAAAATATGTTGCTGGCAAAATGCAAACTGCCAGT 1740
 Qy 1737 AGCAAGAGAGCTCGAAACCGATGTCAGTACTGTCGATTTTTCAGAAAGTGTCTCAGTGTG 1796
 Db 1741 AGCAAGAGAGCTCGAAACCGATGTCAGTACTGTCGATTTTTCAGAAAGTGTCTCAGTGTG 1800
 Qy 1797 AATGTTAAAGAGTGTTCGTAAGATGTCGTAAGAGGAGAGAGTGTGCTGCTTC 1856
 Db 1801 AATGTTAAAGAGTGTTCGTAAGATGTCGTAAGAGGAGAGAGTGTGCTGCTTC 1860
 Qy 1857 CAACCAAGAGCCGCTTACAGAGGAGCTTCTCAGCCCTCTCCACCTTCTCCCTCAAT 1916
 Db 1861 CAACCAAGAGCCGCTTACAGAGGAGCTTCTCAGCCCTCTCCACCTTCTCCCTCAAT 1920
 Qy 1917 CTGCAATGATGATGCTTGTCCGAGCTTTTAAAGAGCTTAAACACCCAGAGATCTTGATTA 1976
 Db 1921 CTGCAATGATGATGCTTGTCCGAGCTTTTAAAGAGCTTAAACACCCAGAGATCTTGATTA 1980
 Qy 1977 TTCCAGATATGCTCCACTGACCGAGCTGCTGAGGAGAGTGTGCAACA 2036
 Db 1981 TTCCAGATATGCTCCACTGACCGAGCTGCTGAGGAGAGTGTGCAACA 2040
 Qy 2037 ATTTCTAACACCTCTGACAGCTTCCATTTGATGATTCAGAAAGCTGGGCGAGAAAGATTCC 2096
 Db 2041 ATTTCTAACACCTCTGACAGCTTCCATTTGATGATTCAGAAAGCTGGGCGAGAAAGATTCC 2100
 Qy 2097 GGGATTTATGATCTCCCAAGAGATCAGACATTACTTATTAATCAGCTTTTTCGA 2156
 Db 2101 GGGATTTATGATCTCCCAAGAGATCAGACATTACTTATTAATCAGCTTTTTCGA 2160
 Qy 2157 GCTGTTGCTCAGACTTTCATCAGGTCAAAACCTGCTGAAGATAAGTTTGTCTG 2216
 Db 2161 GCTGTTGCTCAGACTTTCATCAGGTCAAAACCTGCTGAAGATAAGTTTGTCTG 2220
 Qy 2217 CAATGAGCTTGTCTGCACTCAGTTCAGTTCCTGTTGATTTGGGAGTGGCTGACTC 2276
 Db 2221 CAATGAGCTTGTCTGCACTCAGTTCAGTTCCTGTTGATTTGGGAGTGGCTGACTC 2280
 Qy 2277 TATTAAGACTTTCCTTAAATTTGAGAGCTGAACTTGAATATCCAGGCTTACGCTG 2336
 Db 2281 TATTAAGACTTTCCTTAAATTTGAGAGCTGAACTTGAATATCCAGGCTTACGCTG 2340
 Qy 2337 CTGTCAGCACTGAGCATGATCACAGAAAGATGGTTTAAAGAACCAAGAGAGTCCA 2396
 Db 2341 CTGTCAGCACTGAGCATGATCACAGAAAGATGGTTTAAAGAACCAAGAGAGTCCA 2400

239 CGCCGAGTGAACGCGCCCTCCCGGCTCACTTTGCAACGCTGAGGTCGCGCAGTGGC 298
 241 CGCCGAGTGAACGCGCCCTCCCGGCTCACTTTGCAACGCTGAGGTCGCGCAGTGGC 300
 299 CGTGAGGTGGAAACAGCGGGCGGATCTCCCGGCTGCTGAGGTCGCGCAGGAGCGCC 358
 301 CGTGAGGTGGAAACAGCGGGCGGATCTCCCGGCTGCTGAGGTCGCGCAGGAGCGCC 360
 359 CGCGAACTCTCGGCTGCTGCTCCATGAGTGGGATCGGAGATCCCGCAGGCG 418
 361 CGCGAACTCTCGGCTGCTGCTCCATGAGTGGGATCGGAGATCCCGCAGGCG 420
 419 CTCACGCTCCCGGAGCGGCTGGGCTGTACACCGCAGGCTTCCGCGGACAGAGCTGT 478
 421 CTCACGCTCCCGGAGCGGCTGGGCTGTACACCGCAGGCTTCCGCGGACAGAGCTGT 480
 479 GACTCCCGCCAGTTCAGATTTCCGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC 538
 481 GACTCCCGCCAGTTCAGATTTCCGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC 540
 539 CGCCACAGCACTCAAGCCCACTCGGGAAGAGGAGCGCGGCAAGCGCGGCTGAGC 598
 541 CGCCACAGCACTCAAGCCCACTCGGGAAGAGGAGCGCGGCAAGCGCGGCTGAGC 600
 599 CTGGACCTTAGCGGTGCGGGCAGCACTGCGGGGCTTCCGCTCGCGGAGCTCCGCTC 658
 601 CTGGACCTTAGCGGTGCGGGCAGCACTGCGGGGCTTCCGCTCGCGGAGCTCCGCTC 660
 659 CTCCTACCTCTCAGGCTCGGTGGAGAGACCCCGAGCCCACTTCAAGCGGCAAGAT 718
 661 CTCCTACCTCTCAGGCTCGGTGGAGAGACCCCGAGCCCACTTCAAGCGGCAAGAT 720
 719 ACCTCCAGATATGCGGTCCAGGCTCAAGCCCACTATAGCCCTTCCCTCCAGGTTCCAGTT 778
 721 ACCTCCAGATATGCGGTCCAGGCTCAAGCCCACTATAGCCCTTCCCTCCAGGTTCCAGTT 780
 779 ATCGGGCGGACATACAGCTCGGAATACACCGGAGATCAAGAACCCGCACTACCA 838
 781 ATCGGGCGGACATACAGCTCGGAATACACCGGAGATCAAGAACCCGCACTACCA 840
 839 AGCTGACATGGAACCTTGGGAGCACTGAGATACGGCTACGACACAGCTCCCTGCCCA 898
 841 AGCTGACATGGAACCTTGGGAGCACTGAGATACGGCTACGACACAGCTCCCTGCCCA 900
 899 GCATCAGTACTTCTGAGGAGGCTACTCGAGCACTACGAACTCAAGGCTTCTCGGCTGT 958
 901 GCATCAGTACTTCTGAGGAGGCTACTCGAGCACTACGAACTCAAGGCTTCTCGGCTGT 960
 959 ACCAAATGCAAGCGCCCTTGATCAAGTGGAGAGGGGCGGCGGCGGCTACCATCACC 1018
 961 ACCAAATGCAAGCGCCCTTGATCAAGTGGAGAGGGGCGGCGGCGGCTACCATCACC 1020
 1019 ATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1078
 1021 NNN 1080
 1079 TTCCTCCAGCTCCAGCCCGGAGGAGGAGTGTGCTCCAGCACTCAATGTACTTCAAGC 1138
 1081 TTCCTCCAGCTCCAGCCCGGAGGAGGAGTGTGCTCCAGCACTCAATGTACTTCAAGC 1140
 1139 AGTCCCAACGTCACCCCAACGAGCGGCTTCCCGGCGGAGGCGGCGGCTTATGGG 1198
 1141 AGTCCCAACGTCACCCCAACGAGCGGCTTCCCGGCGGAGGCGGCGGCTTATGGG 1200
 1199 ACAGGAGCACTGCTCCGCGGCTGCGGCTGATCGACCGCGGCGGCTGCTGAGACCGCGCA 1258
 1201 ACAGGAGCACTGCTCCGCGGCTGCGGCTGATCGACCGCGGCGGCTGCTGAGACCGCGCA 1260
 1259 TGAAGGCGGTCCCAACGAGTGGCGGCGGCTTCCGCGCTTCCCACTTCAAGCCCTCGC 1318
 1261 TGAAGGCGGTCCCAACGAGTGGCGGCGGCTTCCGCGCTTCCCACTTCAAGCCCTCGC 1320
 1319 CGCCGATCCCCCGCGCCAGCCCGCGGCGGCGGCGGCGGCTACGACCCGAGCGG 1378

1321 CGCCGATCCCCCGCGCCAGCCCGCGGCGGCGGCGGCGGCTACGACCCGAGCGG 1380
 1379 CGCCTCGCGGCTCAGCCTGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 1438
 1381 CGCCTCGCGGCTCAGCCTGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 1440
 1439 CGCTTGAGAGCCACCGGTAGCGGCTCGCTGCGCAAGAGGGCGGCGGCGGCTTCC 1498
 1441 CGCTTGAGGGCCACCGGTAGCGGCTCGCTGCGCAAGAGGGCGGCGGCGGCTTCC 1500
 1499 CGCCTTGCGGCTCAGCCTCCCTCAGCGCTCAGCGCTGCGGCGGAGAGTCCAGGC 1558
 1501 CGCCTTGCGGCTCAGCCTCCCTCAGCGCTCAGCGCTGCGGCGGAGAGTCCAGGC 1560
 1559 TGCCTGCGCGCGCCAGCAGGAG-CTGCTGCTGCGGAGGCGCACTGCTGCGGCGG 1617
 1561 TGCCTGCGCGCGCCAGCAGGAGTCTGCTGCTGCGGAGGCGCACTGCTGCGGCGG 1620
 1618 GACAAACCGCGC-CTGCGCAGCACTACGCGGTGCGAACTGCGAGGGCTGCAAGGCTTTT 1676
 1621 GACAAACCGCGC-CTGCGCAGCACTACGCGGTGCGAACTGCGAGGGCTGCAAGGCTTTT 1680
 1677 CAAGAGAACAGTGCAGAAATGCAAAATATGTTGCTGCGCAATAAATACTGCCAGT 1736
 1681 CAAGAGAACAGTGCAGAAATGCAAAATATGTTGCTGCGCAATAAATACTGCCAGT 1740
 1737 AGACAAGAGAG-TCGAAACCGGATGTCAGTACTGTCGATTCAGAAAGTGTCTCAGTGTG 1795
 1741 AGACAAGAGAGTTCGAAACCGGATGTCAGTACTGTCGATTCAGAAAGTGTCTCAGTGTG 1800
 1796 GAATGGTAAAGAGTGTGCTGCTGAGATGTCGAAAGGAGGAGAGTGTCTGCTT 1855
 1801 GAATGGTAAAGAGTGTGCTGCTGAGATGTCGAAAGGAGGAGAGTGTCTGCTT 1860
 1856 CCAACCAAGAGAGGCGCACTTACACAGGAACTTCTCAGCGCTTCTCCTCTCCAA 1915
 1861 CCAACCAAGAGAGGCGCACTTACACAGGAACTTCTCAGCGCTTCTCCTCTCCAA 1920
 1916 TCTGATGATGAATGCGCTTGTGCTGAGCTTTAAAGACTCAACACCGGAGATCTTGT 1975
 1921 TCTGATGATGAATGCGCTTGTGCTGAGCTTTAAAGACTCAACACCGGAGATCTTGT 1980
 1976 ATTCAGATGATGCTCCAGCTGCTGAGGAGGAGGAGTGTGCTGAGATGCTGCAAC 2035
 1981 ATTCAGATGATGCTCCAGCTGCTGAGGAGGAGGAGTGTGCTGAGATGCTGCAAC 2040
 2036 AATTCACAACTCTGAGAGCTCCATGATGATCCAGAGCTGGGAGGAGGATTC 2095
 2041 AATTCACAACTCTGAGAGCTCCATGATGATCCAGAGCTGGGAGGAGGATTC 2100
 2096 CGGATTTTACTGATCTCCCAAGAGATCAGATTTTAAAGTGTGATTCAGGCTTTTGG 2155
 2101 CGGATTTTACTGATCTCCCAAGAGATCAGATTTTAAAGTGTGATTCAGGCTTTTGG 2160
 2156 AGCTGTTTGTCTCAGACTTTCCATCAGGTCAAACTGCTGAGATGATGTTGTCT 2215
 2161 AGCTGTTTGTCTCAGACTTTCCATCAGGTCAAACTGCTGAGATGATGTTGTCT 2220
 2216 GCAATGGAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2275
 2221 GCAATGGAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
 2276 CTATTAAGAGCTTTTCTTAAATTTGAGAGGCTGAACTTGAATCAAGCTTTAGCT 2335
 2281 CTATTAAGAGCTTTTCTTAAATTTGAGAGGCTGAACTTGAATCAAGCTTTAGCT 2340
 2336 GCTGTCAGCACTGAGCATGATCAAGAAAGATGATGATGATGATGATGATGATGATGAT 2395
 2341 GCTGTCAGCACTGAGCATGATCAAGAAAGATGATGATGATGATGATGATGATGATGAT 2400
 2396 AAGAGCTATGCAACAGATCAAGAGCTTTAAAGAGACCAACAGAGTAAAGGAGAGGCTC 2455

Db 2401 AAGAGCTATGCAACAAGATCACAAGCAGTTTAAAGACCAACAGAGTAAGGGACAGGCTC 2460
QY 2456 TGGAGCCACCGAGTCAAGGTCTCTGGTGCCCTGTGAGAACTGAGAGAAATCTGCACCC 2515
Db 2461 TGGAGCCACCGAGTCAAGGTCTCTGGTGCCCTGTGAGAACTGAGAGAAATCTGCACCC 2520
QY 2516 TGGGCTCCAGCCGATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTCCATCA 2575
Db 2521 TGGGCTCCAGCCGATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTCCATCA 2580
QY 2576 TTGACAAGCTCTTCTGGACACCTTCTTAATCAGGAGCAGTGGAGCAGTGAAGCTG 2635
Db 2581 TTGACAAGCTCTTCTGGACACCTTCTTAATCAGGAGCAGTGGAGCAGTGAAGCTG 2640
QY 2636 CTTCTCTCTAGCAGCTCTTCTGCTACGACGAAAGGATAGTGTGGAAACCTATCAT 2695
Db 2641 CTTCTCTCTAGCAGCTCTTCTGCTACGACGAAAGGATAGTGTGGAAACCTATCAT 2700
QY 2696 TCTGTCTCTCTTAAAGGAAAGCAGCTCTCTGTAGAAGCAAGACTTCTTTT 2755
Db 2701 TCTGTCTCTCTTAAAGGAAAGCAGCTCTCTGTAGAAGCAAGACTTCTTTT 2760
QY 2756 CTGGCTCTTCTTCAACCTTAAAGCCAGAAACTTGCAGAGTATGTGTGGGCTGT 2815
Db 2761 CTGGCTCTTCTTCAACCTTAAAGCCAGAAACTTGCAGAGTATGTGTGGGCTGT 2820
QY 2816 GTTTTATATTAGGCAATTGGGGATGGGGTGGAGGGGGTTATAGTTCATGAGGGTTTC 2875
Db 2821 GTTTTATATTAGGCAATTGGGGATGGGGTGGAGGGGGTTATAGTTCATGAGGGTTTC 2880
QY 2876 TAAGAAATTGCTTAACAAGCACTTTTGGACAAATGCTATCCAGCAGCAAAAAGGATA 2935
Db 2881 TAAGAAATTGCTTAACAAGCACTTTTGGACAAATGCTATCCAGCAGCAAAAAGGATA 2940
QY 2936 ATATACTGTTTAAACCTTTCTGGGAATCCAATTATAGTGTCTTGTATTTAAAAA 2995
Db 2941 ATATACTGTTTAAACCTTTCTGGGAATCCAATTATAGTGTCTTGTATTTAAAAA 3000
QY 2996 CAAGAACAGCAAGGGTGTTCGCCAGGGTAGATGTCTTAAAGATTGGTCCCTTGAA 3055
Db 3001 CAAGAACAGCAAGGGTGTTCGCCAGGGTAGATGTCTTAAAGATTGGTCCCTTGAA 3060
QY 3056 AATATGCTCTCTATCAAGGTACGTATGTGTGCAAAAGGCAAACTTCTTTT 3115
Db 3061 AATATGCTCTCTATCAAGGTACGTATGTGTGCAAAAGGCAAACTTCTTTT 3120
QY 3116 ATTTCTCTCTCTTATTTTAAACAAATGGTGAAGATGGAGATTACCTTACAAATCAGA 3175
Db 3121 ATTTCTCTCTCTTATTTTAAACAAATGGTGAAGATGGAGATTACCTTACAAATCAGA 3180
QY 3176 CATGGCAAAACAATAATGGCTGTTTGTCCATPAAACAAAGTGCATTTTAAAGTGCTG 3235
Db 3181 CATGGCAAAACAATAATGGCTGTTTGTCCATPAAACAAAGTGCATTTTAAAGTGCTG 3240
QY 3236 TCTTACTAGTCTGTGTTTATTAATCTCTCTTTTATCTATATGGAATAAAGAGGCGAG 3295
Db 3241 TCTTACTAGTCTGTGTTTATTAATCTCTCTTTTATCTATATGGAATAAAGAGGCGAG 3300
QY 3296 TCATGTTAGCAATGACAGCTTAATATCCCTAGCAGGCTGTGTTTCACTTCCCTGTGCG 3355
Db 3301 TCATGTTAGCAATGACAGCTTAATATCCCTAGCAGGCTGTGTTTCACTTCCCTGTGCG 3360
QY 3356 ATCCCTTCTAGGATATGGCCCATCCAAAGCTTTTAGGCCATTTCTGATGGAACCATGTC 3415
Db 3361 ATCCCTTCTAGGATATGGCCCATCCAAAGCTTTTAGGCCATTTCTGATGGAACCATGTC 3420
QY 3416 CTGCTGCTGCTGCTGCTATCTGAAAGTGGATCAGATTAATACTGATTAACATGTAA 3475
Db 3421 CTGCTGCTGCTGCTGCTATCTGAAAGTGGATCAGATTAATACTGATTAACATGTAA 3480
QY 3476 CTGTTTTGGTGTGTTCTATCAACCCACAGAGTTCCCTTAACTTGTCTTCACTATAGT 3535
Db 3481 CTGTTTTGGTGTGTTCTATCAACCCACAGAGTTCCCTTAACTTGTCTTCACTATAGT 3540

QY 3536 AACTGACTGGTATATTCATTTCAGAGCGCCATTAAGTCAAGTTCAGTATTTGATCCCTAGAT 3595
Db 3541 AACTGACTGGTATATTCATTTCAGAGCGCCATTAAGTCAAGTTCAGTATTTGATCCCTAGAT 3600
QY 3596 AAGAACATCAAAATCAGCAGGAACCTGGTCAATACAGGTTAAGCACCAAGGCAATAAGGAT 3655
Db 3601 AAGAACATCAAAATCAGCAGGAACCTGGTCAATACAGGTTAAGCACCAAGGCAATAAGGAT 3660
QY 3656 TTTTATAGATATAAATTTAATTTTGTATTTGGTTAAGGAGACAATTTTGGAGAGCAAGCA 3715
Db 3661 TTTTATAGATATAAATTTAATTTTGTATTTGGTTAAGGAGACAATTTTGGAGAGCAAGCA 3720
QY 3716 AA---TCTTTTAAATAATAGTATGTAATCTAGAAAGATTTTAAATAATAGTAT 3772
Db 3721 AATCTTCTTTTAAATAATAGTATGTAATCTAGAAAGATTTTAAATAATAGTAT 3780
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794
Db 3781 GAGTGTGAGTACTAGGAAGGAT 3802

RESULT 8

US-09-873-367C-229
; Sequence 229, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 229
; LENGTH: 4977
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-229

Query Match 75.8%; Score 2875; DB 10; Length 4977;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;

QY 733 CCTGTGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGG--CGCAGA 790
Db 94 CTGTGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGTCGTGAGA 153
QY 791 CATACAGCTCGGAATACACACCGAGATCATGAACCCCGACATACACCAAGCTGACCATGG 850
Db 154 CATACAGCTCGGAATACACACCGAGATCATGAACCCCGACATACACCAAGCTGACCATGG 213
QY 851 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACACGCTCCCTGCCAGCATCAGTACCT 910
Db 214 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACCACTGCTCCCTGCCAGCATCAGTACCT 273
QY 911 TCGTGGAGGGCTACTCGAGCAACTACGAACTCAAGCCTTCTTCCGTTGTTACCAATGCGAGC 970
Db 274 TTGTGGAGGGCTACTCGAGCAACTACGAACTCAAGCCTTCTTCCGTTGTTACCAATGCGAGC 333

QY	971	GGCCCTTGATCAAGTGGAGGAGGGCGGGCGCCAGCTACCATCACCAACACC	1030	QY	2051	TGACAGCTCCATTGATGATATCCAGAGCTGGCGAGAAAGATTCGGGATTTACTGATC	2110
DB	334	GGCCCTTGATCAAGTGGAGGAGGGCGGGCGCCAGCTACCATCACCAACACC	393	DB	1414	TGACAGCTCCATTGATGATATCCAGAGCTGGCGAGAAAGATTCGGGATTTACTGATC	1473
QY	1031	AC	1090	QY	2111	TCGCCAAAGAAATCAGACATTAATTAATTAATTAATTAATTAATTAATTAATTA	2170
DB	394	AC	453	DB	1474	TCGCCAAAGAAATCAGACATTAATTAATTAATTAATTAATTAATTAATTAATTA	1533
QY	1091	CCAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1150	QY	2171	GACTTTCATCAGGTCAACACACACACACACACACACACACACACACACACAC	2230
DB	454	CCAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	513	DB	1534	GACTTTCATCAGGTCAACACACACACACACACACACACACACACACACACAC	1593
QY	1151	CCACCCCCAC	1210	QY	2231	TGCATCAGCTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTATTAAGA	2290
DB	514	CCACCCCCAC	573	DB	1594	TGCATCAGCTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTATTAAGA	1653
QY	1211	CCTCGGCGCCGCTGCATCGACACCGGCGCGCTGCTGACCCGCGATGAAGCGG	1270	QY	2291	CCTTAATTTGCGAGGCTGAACTTCATATCAAGCCTTAGCCTGCTGACGACTGA	2350
DB	574	CCTCGGCGCCGCTGCATCGACACCGGCGCGCTGCTGACCCGCGATGAAGCGG	633	DB	1654	CCTTAATTTGCGAGGCTGAACTTCATATCAAGCCTTAGCCTGCTGACGACTGA	1713
QY	1271	CCACGCTGGCGCGCGCGCTTCCTCGCTCTTCCACTTCAAGCCCTCGCGCCAT	1330	QY	2351	GCATGATCACAAGAAACATGGGTTAAAGAACCAAAGAGAGTCGAAGAGCTAT	2410
DB	634	CCACGCTGGCGCGCGCGCTTCCTCGCTCTTCCACTTCAAGCCCTCGCGCCAT	693	DB	1714	GCATGATCACAAGAAACATGGGTTAAAGAACCAAAGAGAGTCGAAGAGCTAT	1773
QY	1331	CCGCGCCAGCCG	1390	QY	2411	AGATCACAAGCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCGGAGCCAC	2470
DB	694	CCGCGCCAGCCG	753	DB	1774	AGATCACAAGCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCGGAGCCAC	1833
QY	1391	TCAGCTTCGCGCTGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1450	QY	2471	CCAGGTCTCGGTGGCGCTCGTAGAATCTGAGGAAGATCTGCACCTGGGCGCT	2530
DB	754	TCAGCTTCGCGCTGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	813	DB	1834	CCAGGTCTCGGTGGCGCTCGTAGAATCTGAGGAAGATCTGCACCTGGGCGCT	1893
QY	1451	ACCCTGACGGCTGCGCTGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG	1510	QY	2531	TCCTCTACCTGAACTGGAAGATTTGGTGTCTCCACTTCATCATTTGCAAGCT	2590
DB	814	ACCCTGACGGCTGCGCTGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG	873	DB	1894	TCCTCTACCTGAACTGGAAGATTTGGTGTCTCCACTTCATCATTTGCAAGCT	1953
QY	1511	TCAGCTTCGCGCTGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1570	QY	2591	TGACACCTTACCTTTCTAATCAGAGCGTGGAGCAGTGGAGCTCTCTCTCTAG	2650
DB	874	TCAGCTTCGCGCTGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	933	DB	1954	TGACACCTTACCTTTCTAATCAGAGCGTGGAGCAGTGGAGCTCTCTCTCTAG	2013
QY	1571	CCAGCAGAGCTCGTCTGCTGGCGAGGCGACGTGTGCGTGGGCGGCAACGCG	1630	QY	2651	CCTGCTCTCAGCAGCAAGGATAGGTTTGGAAACCTATCATTTCTGCTCTCTTA	2710
DB	934	CCAGCAGAGCTCGTCTGCTGGCGAGGCGACGTGTGCGTGGGCGGCAACGCG	993	DB	2014	CCTGCTCTCAGCAGCAAGGATAGGTTTGGAAACCTATCATTTCTGCTCTCTTA	2073
QY	1631	GCCAGCAGTACGGCGTGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG	1690	QY	2711	AGAGAAAGCAGCTCTGTAGAAAGCAAGACCTTTCTTTTCTGGCTCTTTCTCT	2770
DB	994	GCCAGCAGTACGGCGTGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG	1053	DB	2074	AGAGAAAGCAGCTCTGTAGAAAGCAAGACCTTTCTTTTCTGGCTCTTTCTCT	2133
QY	1691	AGAAAAATGCAAAATATGTTTGGCGAAATAAAACTGCCAGTAGACAAGAGCG	1750	QY	2771	ACAACTTAAAGCCAGAAACCTTGACAGATTTGTTGGGGTGTGTTTATATTAG	2830
DB	1054	AGAAAAATGCAAAATATGTTTGGCGAAATAAAACTGCCAGTAGACAAGAGCG	1113	DB	2134	ACAACTTAAAGCCAGAAACCTTGACAGATTTGTTGGGGTGTGTTTATATTAG	2193
QY	1751	GAACCCGATGTCAGTCTGATTCAGAGTGTCTCAGTGTGGATGGTAAAGAG	1810	QY	2831	ATTGGGGATGGGGTGGGAGGGGTTATAGTTTATAGGGTTCCTAAGAAATGCT	2890
DB	1114	GAACCCGATGTCAGTCTGATTCAGAGTGTCTCAGTGTGGATGGTAAAGAG	1173	DB	2194	ATTGGGGATGGGGTGGGAGGGGTTATAGTTTATAGGGTTCCTAAGAAATGCT	2253
QY	1811	TTGTCCTGACAGTGTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1870	QY	2891	AAAGCAGCTTTGGCAATGCTATCCAGCGGAAAAAAGAGATATTAATCTGTTTAA	2950
DB	1174	TTGTCCTGACAGTGTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1233	DB	2254	AAAGCAGCTTTGGCAATGCTATCCAGCGGAAAAAAGAGATATTAATCTGTTTAA	2313
QY	1871	CATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCTCTCCAAATCTG	1930	QY	2951	AACTCTTCTCGGGGAATCCAATTAATAGTTGCTTTTATTAATAAACAAGAAC	3010
DB	1234	CATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCTCTCCAAATCTG	1293	DB	2314	AACTCTTCTCGGGGAATCCAATTAATAGTTGCTTTTATTAATAAACAAGAAC	2373
QY	1931	CCCTGTCGAGCTTTAAGACACTCAACCCAGAGATCTTGATTTTCCAGATAC	1990	QY	3011	GTGTTTGGCAGGGTAGGATGCTTAAAGATTTGGTCCCTTGAATAATGCTCTGTA	3070
DB	1294	CTCTTGTCCGAGCTTTAAGACACTCAACCCAGAGATCTTGATTTTCCAGATAC	1353	DB	2374	GTGTTTGGCAGGGTAGGATGCTTAAAGATTTGGTCCCTTGAATAATGCTCTGTA	2433
QY	1991	CCACTGACAGGCTGCTGAGGCGACAGATGCTGAGCATGTGCAACAAATTTTAC	2050	QY	3071	TCAGAGTACGATGTTGGTGGCAAAACAAGCAGAACTTCTTTTAAATTTCTTCT	3130
DB	1354	CCACTGACAGGCTGCTGAGGCGACAGATGCTGAGCATGTGCAACAAATTTTAC	1413	DB	2434	TCAGAGTACGATGTTGGTGGCAAAACAAGCAGAACTTCTTTTAAATTTCTTCT	2493
				QY	3131	TATTTTAAACAAATGGTGAAGATGGAGGATTAACCTACAAATCAGACATGGCA	3190

```
Db 2494 |||||||TATTTTAAATGGTGAAGATGGAGATTAATCTAC-AATCAGATGGCAAAAT- 2551
Qy 3191 ATGGCTGTTTGGCTTCCATAAACAAGTGCAATTTTTTAAAGTGCTGTCTTACTAAGTCTTG 3250
Db 2552 ATGGCTGTTTGGCTTC--ATAACAGCTGCAATTTTTTAAAGTGCTGTCTTACTAAGTCTTG 2608
Qy 3251 TTTATTAATCTCTCTTTATCTATATATGGAATTAATAAGAGGAGGAGTCATGTTAGCAATG 3310
Db 2609 T--TATATCTCTTATCTATATGCGCGAAATATAAAGAGGAGGAGTCATGTTAGCAATG 2665
Qy 3311 ACACGTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCCTGTGATCCCTTCTGAGGTA 3370
Db 2666 ACACGTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCCTGTGATCCCTTCTGAGGTA 2725
Qy 3371 TGGCCCATCAAGACTTTTAGGCCATTTCTGTGATGGAACAGATCCCTGCGCTGACTGTCC 3430
Db 2726 TGGCCCATCAAGACTTTTAGGCCATTTCTGTGATGGAACAGATCCCTGCGCTGACTGTCC 2785
Qy 3431 AGCTATCCTGAAAGTGGATCAGATTATAAAGTGGATTAATGATGATGATGATGATGATGAT 3490
Db 2786 AGCTATCCTGAAAGTGGATCAGATTATAAAGTGGATTAATGATGATGATGATGATGATGAT 2845
Qy 3491 TCTATCAACCCCAACAGAGTTCCTAAACTTCTTCACTTATAGTAAGTAACTGACTGTATAT 3550
Db 2846 TCTATCAACCCCAACAGAGTTCCTAAACTTCTTCACTTATAGTAAGTAACTGACTGTATAT 2904
Qy 3551 TCATTCAAGAGGCCATAAGTCAAGTGGATTTTATGATGATGATGATGATGATGATGATGAT 3610
Db 2905 TCATTCAAGAGGCCATAAGTCAAGTGGATTTTATGATGATGATGATGATGATGATGATGAT 2959
Qy 3611 AGCAGAACTGTCATACAGGTAAGCAACAGGCAATAAGATTTTATGATGATGATGATGATGAT 3670
Db 2960 TCAGAGGACTGTCATACAGGTAAGCAACAGGCAATAAGATTTTATGATGATGATGATGATGAT 3019
Qy 3671 TTAATTTTGTATTGTTTAAAGAGACAAATTTTGGAGAGCAAGCAAA---TCTTTTAA 3727
Db 3020 TTAATTTTGTATTGTTTAAAGAGACAAATTTTGGAGAGCAAGCAAACTCTCTTTTAA 3079
Qy 3728 AAATAGTATGAATGTGAATACTAGAAAGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 3787
Db 3080 AAATAGTATGAATGTGAATACTAGAAAGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 3139
Qy 3788 GAAGGAT 3794
Db 3140 GAAGGAT 3146
```

RESULT 9

```
US-10-755-889-63
; Sequence 63, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; PRIOR FILING DATE: 2004-01-13
; PRIOR FILING DATE: 2003-01-14
; PRIOR FILING DATE: 2003-01-14
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63
; LENGTH: 4977
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-63
```

Query Match 75.8%; Score 2875; DB 18; Length 4977;
Best Local Similarity 98.1%; Pred. No. 0;

```
Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;
Qy 733 CCTCGCTCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATCGG--CGCAGA 790
Db 94 CTTGTGCTCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATCGGTCGTCAGA 153
Qy 791 CATACAGCTCGGAATACACACGAGATCATGAACCCGACTACACACAGCTGACCATGG 850
Db 154 CATACAGCTCGGAATACACACGAGATCATGAACCCGACTACACACAGCTGACCATGG 213
Qy 851 ACTTTGGCAGCACTGAGATCAGGGCTACAGCACCACAGCTCCCTGCCAGCATCAGTACCT 910
Db 214 ACTTTGGCAGCACTGAGATCAGGGCTACAGCACCACAGCTCCCTGCCAGCATCAGTACCT 273
Qy 911 TGTGTAGAGGCTTACTCGAGCAATAGAACTCAAGCTTCTGCGTGTACCAAAATGACG 970
Db 274 TTTGTGAGGCTTACTCGAGCAATAGAACTCAAGCTTCTGCGTGTACCAAAATGACG 333
Qy 971 GGCCTTTGATCAAGTGGAGGAGGGCGGCCAGCTACCATCACCATCAACCACACC 1030
Db 334 GGCCTTTGATCAAGTGGAGGAGGGCGGCCAGCTACCATCACCATCAACCACACC 393
Qy 1031 ACCACACACACACACACCATCAACAGCAGCAGCATCAGCAGGATCCATTTCTTCCAGCCT 1090
Db 394 ACCACACACACACACACCATCAACAGCAGCAGCATCAGCAGGATCCATTTCTTCCAGCCT 453
Qy 1091 CCAGCCCGAGAGCAGAGGTGCTGCCAGCACTTCCATGTACTTCAAGAGCTGCCACCGT 1150
Db 454 CCAGCCCGAGAGCAGAGGTGCTGCCAGCACTTCCATGTACTTCAAGAGCTGCCACCGT 513
Qy 1151 CCACCCCAACAGCGCGGCTTCCCTCCAGCGCGGGGGGCTTATGGGACGAGGCACTGC 1210
Db 514 CCACCCCAACAGCGCGGCTTCCCTCCAGCGCGGGGGGCTTATGGGACGAGGCACTGC 573
Qy 1211 CTTGCGCGCGCGCTGTCACACCGCGCGCTGCTGACCCGCGATGAAGCGCGTCC 1270
Db 574 CTTGCGCGCGCGCTGTCATCGACCGCGCGCTGCTGACCCGCGATGAAGCGCGTCC 633
Qy 1271 CCACGCTGCGCGCGCGCTTCCCGCTTCCATTTCAAGGCTTCCGCGCGCATCTCCC 1330
Db 634 CCACGCTGCGCGCGCGCTTCCCGCTTCCATTTCAAGGCTTCCGCGCGCATCTCCC 693
Qy 1331 CCGCGCCACAGCGCGCGCGCGCCACCTCGGCTACACCCGAGCGCGCTGCCGCGC 1390
Db 694 CCGCGCCACAGCGCGCGCGCGCCACCTCGGCTACACCCGAGCGCGCTGCCGCGC 753
Qy 1391 TCAGCTTGCCTGGGAGCGCAGCGCGCGCGAGCAGCGCGCGCTTGAAGGCG 1450
Db 754 TCAGCTTGCCTGGGAGCGCAGCGCGCGCGAGCAGCGCGCGCTTGAAGGCG 813
Qy 1451 ACCCGTACGCGGCTGCGCTGGCCAAAGAGGCGCGCGCGCTTCCGCGCTTCCGCGC 1510
Db 814 ACCCGTACGCGGCTGCGCTGGCCAAAGAGGCGCGCGCGCTTCCGCGCTTCCGCGC 873
Qy 1511 TCAGCGCTTCCCTACCGCGCTCAGCGCTGCTGGGAGAGTCCACGCTGCCGCTGCCGCG 1570
Db 874 TCAGCGCTTCCCTACCGCGCTCAGCGCTGCTGGGAGAGTCCACGCTGCCGCTGCCGCG 933
Qy 1571 CCAGCAGGAGCTGCTGCTGCGAGGCGCAGCTGCTGCTGCGGCGCAACCGCGCTT 1630
Db 934 CCAGCAGGAGCTGCTGCTGCGAGGCGCAGCTGCTGCTGCGGCGCAACCGCGCTT 993
Qy 1631 GCCAGCACTACGCGGTGCGAACTTGCAGGCGTGCAGAGGCTTTTCAAGAGAACAGTGC 1690
Db 994 GCCAGCACTACGCGGTGCGAACTTGCAGGCGTGCAGAGGCTTTTCAAGAGAACAGTGC 1053
Qy 1691 AGAAAAATGCAAAATATGTTTGGCTGGCAAAATAAACTGCCAGTAGAGACAGGCTC 1750
Db 1054 AGAAAAATGCAAAATATGTTTGGCTGGCAAAATAAACTGCCAGTAGAGACAGGCTC 1113
Qy 1751 GAAACCGATGTCTGATCTGATTTTCAAGAGTGTCTGATGTTTGAATGTTAAAGAG 1810
Db 1114 GAAACCGATGTCTGATCTGATTTTCAAGAGTGTCTGATGTTTGAATGTTAAAGAG 1173
```

1811 TTGTCGGTACAGATAGCTGAAAGGAGGAGAGGTGCTGCTGCTTCCAAACCAAGAGGCC 1870
1174 TTGTCGGTACAGATAGCTGAAAGGAGGAGAGGTGCTGCTGCTTCCAAACCAAGAGGCC 1233
1871 CATTACAACAGGAACCTTCTGAGCCCTCTCACCCTTCTCCTCCAACTGCGATGATGAATG 1930
1234 CATTACACAGGAACCTTCTGAGCCCTCTCACCCTTCTCCTCCAACTGCGATGATGAATG 1293
1931 CCGTTGTCGAGCTTAAACAGACTCAACACCCAGAGATCTTGATTAATTCAGATAGTGC 1990
1294 CTCTGTCGAGCTTAAACAGACTCAACACCCAGAGATCTTGATTAATTCAGATAGTGC 1353
1991 CCACTACAGAGCTGCTGAGGACAGAGCTGAGCATGTGCAACAACTTCTACAACTCC 2050
1354 CCACTACAGAGCTGCTGAGGACAGAGCTGAGCATGTGCAACAACTTCTACAACTCC 1413
2051 TGACAGCTCCATTGATGATCCAGAAAGCTGGGCGAAGAAAGATTCCGGGATTTACTGATC 2110
1414 TGACAGCTCCATTGATGATCCAGAAAGCTGGGCGAAGAAAGATTCCGGGATTTACTGATC 1473
2111 TCCCAAGAGATCAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2170
1474 TCCCAAGAGATCAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1533
2171 GACTTTCATCAGCTCAAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGGAATTTGTC 2230
1534 GACTTTCATCAGCTCAAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGGAATTTGTC 1593
2231 TGCATCGACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2290
1594 TGCATCGACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1653
2291 CCTTAAATTTGACAGCTGAACTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2350
1654 CCTTAAATTTGACAGCTGAACTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1713
2351 GCATGATCAGAAAGACATGGGTAAAGAAACCAAGAGAGTTCGAAGAGCTATGCAACA 2410
1714 GCATGATCAGAAAGACATGGGTAAAGAAACCAAGAGAGTTCGAAGAGCTATGCAACA 1773
2411 AGATCAAGAGCTTAAAGACCAACAGAGTAAAGAGAGTTCGAGCCACCCAGT 2470
1774 AGATCAAGAGCTTAAAGACCAACAGAGTAAAGAGAGTTCGAGCCACCCAGT 1833
2471 CCAAGGCTCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2530
1834 CCAAGGCTCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1893
2531 TCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTGACAGCTCTTCC 2590
1894 TCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTGACAGCTCTTCC 1953
2591 TGGACACCTTACCTTCTATCAGAGCAGTGGAGCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2650
1954 TGGACACCTTACCTTCTATCAGAGCAGTGGAGCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2013
2651 CCGTCTGCTACGACGAAAGGATAGGTTTGGAAACCTATCAATTTCTGCTGCTGCTGCTGCTGCT 2710
2014 CCGTCTGCTACGACGAAAGGATAGGTTTGGAAACCTATCAATTTCTGCTGCTGCTGCTGCTGCT 2073
2711 AGAGAAAGAGCTCTCTGTAAGAAAGCAACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2770
2074 AGAGAAAGAGCTCTCTGTAAGAAAGCAACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2133
2771 ACAACCTAAGCCAGAAACTTGCAGAGTATGTTGGGCTTGTGTTTATATTTAGGC 2830
2134 ACAACCTAAGCCAGAAACTTGCAGAGTATGTTGGGCTTGTGTTTATATTTAGGC 2193
2831 ATTGGGGATGGGTGGAGGGGGTTATAGTTCATGAGGGTTTCTTAAGAAATTTGCTAAC 2890
2194 ATTGGGGATGGGTGGAGGGGGTTATAGTTCATGAGGGTTTCTTAAGAAATTTGCTAAC 2253

2891 AAAGCACTTTTGGACAATGCTATCCAGCAGGAGGAAAAAAGGATAATATACTGTTTAA 2950
2254 AAAGCACTTTTGGACAATGCTATCCAGCAGGAGGAAAAAAGGATAATATACTGTTTAA 2313
2951 AACTCTTTTGGGGAATCCAAATATATAGTTGCTTTGATTTTAAAAACAAGAACAGCAAGG 3010
2314 AACTCTTTTGGGGAATCCAAATATATAGTTGCTTTGATTTTAAAAACAAGAACAGCAAGG 2373
3011 GTTGTTCGCGAGGATAGGATGCTCTTAAAGATTGGTCCCTTGAATAATATGCTTCTGTA 3070
2374 GTTGTTCGCGAGGATAGGATGCTCTTAAAGATTGGTCCCTTGAATAATATGCTTCTGTA 2433
3071 TCAAGGTAGCTATGCTGCTGCAAAACAAGCAGAGAACTTCCCTTTTAAATTTCTTCTTCTT 3130
2434 TCAAGGTAGCTATGCTGCTGCAAAACAAGCAGAGAACTTCCCTTTTAAATTTCTTCTTCTT 2493
3131 TATTTTAAACAAATGCTGCAAGATGAGGATTTACCTCAAAATCAGACATGCGCAAAACAATA 3190
2494 TATTTTAAACAAATGCTGCAAGATGAGGATTTACCTCAAAATCAGACATGCGCAAAACAATA 2551
3191 ATGCTGCTTTGCTTCCATAAACAAGTGCATTTTAAAGTGTCTGCTTAAAGTCTTCTTCTTCTT 3250
2552 ATGCTGCTTTGCTTCCATAAACAAGTGCATTTTAAAGTGTCTGCTTAAAGTCTTCTTCTTCTT 2608
3251 TTTTATTAACCTCTCTTTTCTATATGCAATATAAAGAGGAGGAGTCAATGTTAGCAAAATG 3310
2609 T---TATACTCTCTTATCTATATGCGGAAATAAAGAGGAGGAGTCAATGTTAGCAAAATG 2665
3311 ACAGTTTAAATCCTTAGCAGAGGCTGTGTTCACCTTCCCTGTCGATCCCTTCTGAGGTA 3370
2666 ACAGTTTAAATCCTTAGCAGAGGCTGTGTTCACCTTCCCTGTCGATCCCTTCTGAGGTA 2725
3371 TGGCCATCCAAAGACTTTTAGGCCATTTTGAATGCAACCCAGATCCCTGCGCTGACTGTCC 3430
2726 TGGCCATCCAAAGACTTTTAGGCCATTTTGAATGCAACCCAGATCCCTGCGCTGACTGTCC 2785
3431 AGCTATCCTGAAAGTGAATCAGATTATAAACTGGATTACATGTAACCTGTTTGGTGTGT 3490
2786 AGCTATCCTGAAAGTGAATCAGATTATAAACTGGATTACATGTAACCTGTTTGGTGTGT 2845
3491 TCTATCAACCCACAGAGTTCCCTTAACTTGTCTGATTAAGTAACTGACTGCTGATAT 3550
2846 TCTATCAACCCACAGAGTTCCCTTAACTTGTCTGATTAAGTAACTGACTGCTGATAT- 2904
3551 TCATTCAGAGCGCCATAAGTCAAGTTGATGATTTGATCCCTAGATAAGAACATGCAATC 3610
2905 TCATTCAGAGCGCCATAAGTCAAGTTGATGATTTGATCCCTAGATAAGAACATGCAATC- 2959
3611 AGCAGGAACCTGCTCATACAGGTAAGCACAGGGAACAATAAGGATTTTATAGATATAAT 3670
2960 TCAGAGGACTGCTCATACAGGTAAGCACAGGGAACAATAAGGATTTTATAGATATAAT 3019
3671 TTAATTTTGTGTTTAAAGGACAAATTTTGGAGAGCAAGCAAA--TCTTTTAA 3727
3020 TTAATTTTGTGTTTAAAGGACAAATTTTGGAGAGCAAGCAAACTCTTTTAA 3079
3728 AATATAGTATGAATGTAATCTAGAAAAGATTTTAAAAATATAGTATGAGTGTGACTAG 3787
3080 AATATAGTATGAATGTAATCTAGAAAAGATTTTAAAGAAATAGTATGAGTGTGACTAG 3139
3788 GAAGGAT 3794
3140 GAAGGAT 3146

RESULT 10

US-10-641-643-928
; Sequence 928, Application US/10641643
; Publication No. US2004007003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Suan G. Stuart
; Jeffrey J. Seilhamer

Db	1519	CGAGCACTACGGCGTGGAACTCTGCGGGCTGCAAGGGCTTTTTCAGAGAAACAGTGCAC	1578
Qy	1692	GAAAAATGCAAAATATGTTTGCTCGGCAATATAAAACCTGCCAGTAGACAGAGACGTGCG	1751
Db	1579	GAAAAATGCAAAATATGTTTGCTGGCAATATAAAACCTGCCAGTAGACAGAGACGTGCG	1638
Qy	1752	AAACCGATGTCAGTACTGTGCGAATTTTCAGAAAGTGTCTCAGTGTGGAAATGATAAAGAAAGT	1811
Db	1639	AAACCGATGTCAGTACTGTGCGAATTTTCAGAAAGTGTCTCAGTGTGGAAATGATAAAGAAAGT	1698
Qy	1812	TGTCGGTACAGATAGTCTGAAAGGGAGGAGAGGTGCTGTGCGCTTCCAAACCAAAGAGGCC	1871
Db	1699	TGTCCGTACAGATAGTCTGAAAGGGAGGAGAGGTGCTGTGCGCTTCCAAACCAAAGAGGCC	1758
Qy	1872	ATTACAAACGGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAAATCTGCATGATGAATGC	1931
Db	1759	ATTACAAACGGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAAATCTGCATGATGAATGC	1818
Qy	1932	CCTTGTCCGAGCTTTTAAACAGATCTCAACACCCAGAGATCTTGATTAATTCAGATACTGTCC	1991
Db	1819	CCTTGTCCGAGCTTTTAAACAGATCTCAACACCCAGAGATCTTGATTAATTCAGATACTGTCC	1878
Qy	1992	CATGACCCAGGCTGCTGCAGGCACAGATCTGAGCATGTGCAACAATTTCTACAACTCCT	2051
Db	1879	CATGACCCAGGCTGCTGCAGGCACAGATCTGAGCATGTGCAACAATTTCTACAACTCCT	1938
Qy	2052	GACAGCCTCCAATTGATGTATCCAGAAAGCTGGGCAGAAAAAGATTCGGGGATTTTCTGATCT	2111
Db	1939	GACAGCCTCCAATTGATGTATCCAGAAAGCTGGGCAGAAAAAGATTCGGGGATTTTCTGATCT	1998
Qy	2112	CCCCAAAGAAGATCAGACATTAATTATGAATCAGCCCTTTTGGAGCTGTTTGTCTCTCAG	2171
Db	1999	CCCCAAAGAAGATCAGACATTAATTATGAATCAGCCCTTTTGGAGCTGTTTGTCTCTCAG	2058
Qy	2172	ACTTTCATCAGGTCACAAACACTGCTCAGAAATAAGTTTGTGTCTCGAAATGGACTGTGCT	2231
Db	2059	ACTTTCATCAGGTCACAAACACTGCTCAGAAATAAGTTTGTGTCTCGAAATGGACTGTGCT	2118
Qy	2232	GCATCGACTTCAGTCCCTTCGTGGATTTGGGGAGTGGCTCGACTCTATTAAAGACTTTTC	2291
Db	2119	GCATCGACTTCAGTCCCTTCGTGGATTTGGGGAGTGGCTCGACTCTATTAAAGACTTTTC	2178
Qy	2292	CTTAAATTTGCAGAGCTGAAACCTTGATATCCAAAGCCTTAGCCTGCTGTCAACACTGAG	2351
Db	2179	CTTAAATTTGCAGAGCTGAAACCTTGATATCCAAAGCCTTAGCCTGCTGTCAACACTGAG	2238
Qy	2352	CATGATCAAGAAAGACATGCGTTTAAAGAACCAAAGAGATCGAAGAGCTATGCAACAA	2411
Db	2239	CATGATCAAGAAAGACATGCGTTTAAAGAACCAAAGAGATCGAAGAGCTATGCAACAA	2298
Qy	2412	GATCAACAGAGCTTTTAAAGAACCAACAGAGTAAGGGACAGGCTCTGGAGGCCACCGAGTC	2471
Db	2299	GATCAACAGAGCTTTTAAAGAACCAACAGAGTAAGGGACAGGCTCTGGAGGCCACCGAGTC	2358
Qy	2472	CAAGGTCTTGGGTCCCTGCTGAGAACTGAGGAAGATCTGACCCCTGGGGCTCTCAGCGCAT	2531
Db	2359	CAAGGTCTTGGGTCCCTGCTGAGAACTGAGGAAGATCTGACCCCTGGGGCTCTCAGCGCAT	2418
Qy	2532	CTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTCAAGCTCTTCTCT	2591
Db	2419	CTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTCAAGCTCTTCTCT	2478
Qy	2592	GGACACCTTACCTTTCTAATCAGGAGCAGTGGAGAGTGAAGCTGCCTCTCTCTTAGCAC	2651
Db	2479	GGACACCTTACCTTTCTAATCAGGAGCAGTGGAGAGTGAAGCTGCCTCTCTCTTAGCAC	2538
Qy	2652	CTGCTTGTACGACGAAGGGATAGTTTGGAAACCTATCATTTTCTGTCTCTTCTTAA	2711
Db	2539	CTGCTTGTACGACGAAGGGATAGTTTGGAAACCTATCATTTTCTGTCTCTTCTTAA	2598
Qy	2712	GAGGAAAGCAGCTCCTGTAGAAAGCAAGAATTTCTTTTCTTCTGGCTCTTTTCTTTA	2771

Db 2599 GAGGAAAGCAGCTCCTGTGTAAGAACGACATTTCTTTTCTGGCTCTTTTCCTTA 2658

QY 2772 CAACCTAAAGCCAGAAAACCTTGACAGATGTTGTTGGGTTGTTTATATTA 2827

Db 2659 CAACCTAAAGCCAGAAAACCTTGACAGATGTTGTTGGGTTGTTTATATTA 2714

RESULT 11

US-09-853-386-47

; Sequence 47, Application US/09853386

; Patent No. US20020049151A1

GENERAL INFORMATION:

; APPLICANT: Murphy, Evelyn

; APPLICANT: Bresnahan, Barry

; APPLICANT: Conneely, Orla

; APPLICANT: Fitzgerald, Oliver

; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR

; TITLE OF INVENTION: Subfamily of Nuclear Transcription Factors

; FILE REFERENCE: P01972US1

; CURRENT APPLICATION NUMBER: US/09/853,386

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: US 60/203645

; PRIOR FILING DATE: 2000-05-12

; NUMBER OF SEQ ID NOS: 153

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 47

; LENGTH: 4400

; TYPE: DNA

; ORGANISM: Rattus norvegicus

US-09-853-386-47

Query Match 65.9%; Score 2500; DB 9; Length 4400;

Best Local Similarity 83.4%; Pred. No. 0;

Matches 3131; Conservative 0; Mismatches 535; Indels 90; Gaps 22;

QY 46 CGGAGTCTCCTGGCTCCCGCCCCCACCCTCCAGTCTCTGCTCTCTCTCGCTCCCAT 105

Db 1 CGGAGTCTCCTGGCTCCCGCCCCCACCCTCCAGTCTCTGCTCTCTCTCGCTCCCAT 60

QY 106 ACACAGACGCGTCAACCCGCTCCCTCACTCGACACACACAGACGCGGACACAG 165

Db 61 ACACAGACGCGTCAACCCGCTCCCTCACTTGACACACACAGACGCGGCTCACAC 120

QY 166 GCTCGGACACACACT--TGGCTCTCCGCGGCTCACACCCCTCTTGCCCTGAGCCCTT 223

Db 121 GCTCGGACACACACTCGACTCTCTCCGCGGCTCACACCCCTCTCTCTGCGGCCCTC 180

QY 224 GCGGCT----GCAGCGGCGCGCGACGTGAACGCCCTCCCGGGCTCACTTTGCAACG 278

Db 181 GCGGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCCAGGGCTCACTTTGCAACG 240

QY 279 CTGACGGTCCGGACAGTGGCCGTGGAGTGGGAAACAGCGGCGGACATCTCCCCCTGGTC 338

Db 241 CTGACAGACGCGGACAGTGGCCGTGGAGTGGGAAACGTGGGACATCTCTAGCCCTGGTC 300

QY 339 ACAGCCCAAGCCAGACAGCGCCCGGAAACCTCTCGGCTGTGCTCTCCATGAGTCGGGATC 398

Db 301 GCAGCCGGAGA CTGGACG-CTGCGGAAACCTCTCGGCGGCGCTCTCCATGAGTCGGGATC 359

QY 399 GAGCATCCCCACACAGCC---GCTCACCGCTCCGGGAGCGCTGGGCTTGTACACCGC 455

Db 360 GCAGCATCCCCACAGCCGCTGCTCACCGCTCTGGGAGCGCTGGGTTTGTGACACCGC 419

QY 456 AGCCCTTCGGGACACAGCTGTGACTCCCCCCCCAGTGCAGATTTCCGGACAGCTCTCTA 515

Db 420 AGCCCTTCGGGACACAGCTGTGACTCTCCCCCAATCCAGATTTCCGGGCTCGCTCTCTA 479

QY 516 GAACTCGCTCTAAAGACCGGACACAGCACTCAAGGCCACCTGCCGAGAGGCGCAG 575

Db 480 GAACTCGCTCTAAAGACCGGACCTCCACAGAACCCCAAGGCCACCTGCCGAGAGGCGCAG 539

QY 576 CCCGGCAAGCCCGGGCCCTGAGCTCGAACCTTTAGCGGTGCCCGGACGCA-----CTGC 629

Db 540 CCGGACAGCCCGGGCGCTGAGCCTGGACCCTCAACAGAGCGGGCCAGCACAGCGGGCG 599
Qy 630 CGGCGCTTCGCTCGCGGAGCGTCGCTCTCTACACTCTCAGCCTCGCTGGAAGAC 689
Db 600 GGTGCTTCGCTATCCGACGTCGCGGCTCTACACTCTCAGCCTCGCTGGAAGAC 659
Qy 690 CCCAGCCCAACATTCAGCGGCAAGATACCTCCAGATATGCTGCTGCTCAAGCCCA 749
Db 660 CCCAGCCCAACATTCAGCGGCAAGATACCTCCAGATATGCTGCTGCTCAAGCCCA 719
Qy 750 ATATAGCCCTTCGCTTCGAGGTTCCAGTTATGCGGCGCAGACATACAGCTCGGAATACAC 809
Db 720 ATATAGCCCTTCGCTTCGAGGTTCCAGTTATGCGGCGCAGACATACAGCTCGGAATACAC 779
Qy 810 CAGGAGATCATGAACCCCGACTACACCAAGCTGACATGAGACTCTGGAGAGCTAGAT 869
Db 780 CACAGAAATCATGAACCCCGACTATGCCAAGCTGACATGAGACTCTGGAGAGCTAGAT 839
Qy 870 CAGGCTACAGCCACACAGCTCCCTGCGCCAGCATCAGTACTCTGAGAGGCTACTCGAG 929
Db 840 CATGGCCAGCCGACAGAGCTCCCTGCGCCAGCTTCACTACTCTCATGAGAGGCTTACCCAG 899
Qy 930 CAACTACGAATCAAGCCTTCCTGCGTGTACCAAAATGCAG-----CGGCCCTT 977
Db 900 CAGCTCGGAATCAAGCCTTCCTGCGTGTACCAAAATGCCGCTTCTGAGGCTCGGCCCTT 959
Qy 978 GATCAAAAGTGGAGAGGGGGCGGCGAGCTACCATACCATACCAACACACCA 1037
Db 960 GATCAAGATGGAAGAGGGTGGCGAGCATGGCTACCAACACCAACACCATCACCATCA 1019
Qy 1038 CCACCACCACTACACAGCAGCATCAGCAGCCATCCATTCCTCCAGCCTCCAGCCC 1097
Db 1020 TATACCAACCAACAGCAG-----CAGCAGCGCTCCATTCCTCTCGCCCT 1073
Qy 1098 GGAGGACGAGGTGCTGCCAGCAGCTCCATGTACTTCAAGCAGTCCCCACCGTCCACCCC 1157
Db 1074 CGAGGACGAGGTACTGCCAGCAGCTCCATGTACTTCAAGCAGTCTCCGCGCTTACGCC 1133
Qy 1158 CACACGCGCGCTTCGCCCGCAGGGGGGGGTTATGGAGAGGCACTGCGCTCGGC 1217
Db 1134 GACCACTCCAGGCTTCGCCCGCAGGGGGGGGCTGTGGAGCAGCAGCTGCGCTCTGC 1193
Qy 1218 GCCCGCTGCATCGACCCGCGCGCTGTGACCCGCGCATGAAAGCGGTCGCCAGCT 1277
Db 1194 GCCTGCTGCATCGCTCCGAGCAGCTGTGACCCCGCAGATGAAGGCACTGCGCCCAAT 1253
Qy 1278 GGCCTGGCGGCTTCGCCCTTCTTCACTTCAAGCCTTCGCGCGCATCCCGCGCGC 1337
Db 1254 GCGCGTGTGCGGCTTCCGATCTTCTTCAAGCCTTCAACGCAACCTCCCGCGGC 1313
Qy 1338 CAGCCGCGCGGCGGCAACACTCGGCTACGACCGGCGCTGCGCGCTCAGCCT 1397
Db 1314 CAGCCGAGCGCGCGGCAACACTCGGCTATGACCCCGCGCGCAGCTGCGCTCAGTCT 1373
Qy 1398 GCGCTGGAGCGGCGGCGCGCGGCGGCGGCGGCGGCGGCTTGAAGCAGCACCGTA 1457
Db 1374 ACCCTTGGAGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCTTGAAGCAGCACCGTA 1433
Qy 1458 CGGCTGCGGCTGGCCCAAGAGGCGCGCGCTGCGCTTCCCGCTCTCGGCTCAGCGC 1517
Db 1434 CGGCTGCGGCTGGCCCAAGAGGCGCGCGCTTGAAGCAGCACCGCTTCCCGCTCAGCGC 1493
Qy 1518 CTCCTCTACGCGCTCAGCGCTGCTGGCGGAGTCCAGCCTGCGCTGCGCGCGGCGGAG 1577
Db 1494 GTCCCTTACGCGCTCAGCGCTGCTGGGAGAGAGCGGCTTACCATCGCCACCCCAATAG 1553
Qy 1578 GAGCTGCTGCTGGCGGCGGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1637
Db 1554 GAGCTCATCATCGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1613
Qy 1638 CTACGCGCTGCGAACTGCGAGGCTGCAAGGCTTTTTCAGAGAACAGTGCAGAAAAA 1697
Db 1614 CTACGAGTCCGACCTCGAGGCGCTGCAAGGCTTCTTCAGAGAACGCTGCGAGAAAA 1673

Qy 1698 TGCAAAATATGTTTGCCTGGCAAAATAAAACTGCCAGTAGACAAGAGAGCTCGAAACCG 1757
Db 1674 CGCAAAATATGTTTGCCTGGCAAAATAAAACTGCCAGTAGACAAGAGAGCTCGAAATCG 1733
Qy 1758 ATGTCTAGTACTGTGCGAATTTTCAGAGTGTCTCAGTGTGGAATGCTTAAAGAAAGTGTGTCG 1817
Db 1734 ATGTCTAGTACTGTGCGAATTTTCAGAGTGTCTCAGTGTGGAATGCTTAAAGAAAGTGTGTCG 1793
Qy 1818 TACAGTACTGTGAAAGGAGGAGAGGTCTGCTGCTTCCAAACCAAGAGGCCATTTACA 1877
Db 1794 TACAGTACTGTGAAAGGAGGAGAGGTCTGCTGCTTCCAAACCAAGAGGCCATTTACA 1853
Qy 1878 ACAGNAACCTTCTCAGCCCTCTCCACCTTCTCCTCAATCTGCTCATGATGATGATGATGAT 1937
Db 1854 ACAGGAGCCCTCGCAGCCCTCCCACTCTCTCCGATCTGTATGATGATGATGATGATGATGAT 1913
Qy 1938 CCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTTATTCAGATACTGTGCCACTGA 1997
Db 1914 CCGAGCTTTAACAGAGCGAACGCCAGAGACCTTGATTTACTCCAGATACTGTGCCACGA 1973
Qy 1998 CCGGCTGTGCGGCGACAGATGCTGAGCATGTGCAACCAATTTACAACCTCTGACAGC 2057
Db 1974 CCGAGCCACTGCGGCGACAGACGCTGAGCAGCTGCGAGAGTTCTACAACTTCTGACGCG 2033
Qy 2058 CTCCATTGATGTATCCAGAGCTGGGAGAGATTCGGGATTTACTGATCTCCCA 2117
Db 2034 CTCCATTGAGTGTCTCAGAGCTGGGAGAGATTCGGGAGATTCGGGATTCCTGCCCA 2093
Qy 2118 AGAAGATCAGACATTAATTAATCAGCTTTCAGCTTTTGGAGCTGTTGTCTCAGACTTTC 2177
Db 2094 AGAAGATCAGAGCTTACTTATAGATCAGCCTTTTGGAGCTGTTGTCTTAGACTTTC 2153
Qy 2178 CATCAGGTCAACACTGCTGAAAGATAAGTTGTGTTCTGCAATGGAATGCTGCTGCAATG 2237
Db 2154 TATCAGGTCAACACTGCTGAAAGATAAGTTGTGTTCTGCAATGGAATGCTGCTGCAATG 2213
Qy 2238 ACTTCAGTGCCTTCGAGATTTGGGAGTGGGAGTGGCTGAGCTCTATTTAAAGACTTTTCTTAA 2297
Db 2214 ACTTCAGTGCCTTCGCGGATTTGGGAGTGGGCTGAGCTCAATTTAAAGACTTTTCTTAA 2273
Qy 2298 TTTGAGAGCTGAACTTTGATATCAAGCTTTAGCTGCTGCTGAGCAGTGAAGATGAT 2357
Db 2274 TTTGAGAGCTGAACTTTGATATCAAGCTTTAGCTGCTGCTGAGCAGTGAAGATGAT 2333
Qy 2358 CACAGAAAGCATGGGTTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACCAAGATCAC 2417
Db 2334 CACAGAGCAGATGGGTTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACCAAGATCAC 2393
Qy 2418 AAGCAGTTTAAAGACCAACAGATGAAGGACAGGCTCTGGAGCCCAACCGAGTCCAAGGT 2477
Db 2394 AAGCAGTTTAAAGACCAACAGATGAAGGACAGGCTCTGGAGCCCTCAGAGCCCAAGGT 2453
Qy 2478 CTTGGTGCCTTGTGTAAGTGAAGAGATCTGACCTTGGGCTTCCAGGCTCTTCTTA 2537
Db 2454 CTTTGGCAGCTGTTGTAAGTGAAGAGATCTGACCCAGGCTCTCAGGCTATCTTCTA 2513
Qy 2538 CTTGAAGCTGGAAGACTTGGTGTCTCCACTTTCATATGCAAGACTTCTCTGAGACAC 2597
Db 2514 CTTGAAGCTGGAAGACTTGGTGTCTCCACTTTCATATGCAAGACTTCTCTGAGTATC 2573
Qy 2598 CTTACCTTTCTAATCAGAGCAG- TGGAGCAGTGAAGTGGCTCTCTCTAGCAGCTGCT 2656
Db 2574 CTTGCTTTCTGAGCAGGGAAGCTTGGAGAGAGTACTTGTCTGTGTGGCAGCTGCT 2633
Qy 2657 TGTACGCGCAGCAAGGATAGGTTTGGAAACCTTATCATTTCTGTCTCTTCTTAAAGAGGA 2716
Db 2634 ATTAAGTGAAGAAAGATGGGTTGAACACCT- GCCCTCTATCTTCTCTCAGGGGAA 2692
Qy 2717 AAAGCAGCTCTGTAGAAAGCAAGACTTTCTTTTTTTTCTGGCTCTTTTCTTACAAAC 2776
Db 2693 AAAGCAGCTCTCTATAGAAAGCAAGACTTT- TTTTTTCTGSCACTTCTTCTTCTTACAAAC 2751

QY	2777	TAAGCCAGAAACCTGCAGAGTATGTGTGGGGTGTGTTTATATTTAGGCAATTGGG	2836
Db	2752	TAAGCCAGAAACCTGCAGAGTATGTGTGGGGTGTGTTTATATTTAGGCTTTGGT	2811
QY	2837	GGATGGGTGGAGGGGT--TATAGTTTCATGAGGTTTCTTAAGAAATTCGTAAACAAG	2894
Db	2812	GGTGGGCTGGAGGGGTAAATATAGTTCATGAGGCTTCTAAGAAATTCGTGAGGAAG	2871
QY	2895	CACCTTTTGACAATGCTATCCAGCAG-----GAAAAAAGGATAATATACTGTTTT	2948
Db	2872	CACCTTTTGATGATGCTATCCAGCAGTGGGTGGGAGAAAGGATAATAAATCTGTTTT	2931
QY	2949	AAAA--CTCTTTCTGGGAATCCAATATAGTGTCTTCTATTTAAACAGAACAGCC	3006
Db	2932	AAAAACTCTTCCGGGGGAATATGACTATGGTTGCTTTATTTAAATATAGAACAGCC	2991
QY	3007	AAGGGTTG--TTGCCAGGCTAGGATGTGCTTAAAGATTGGTCCCTTGAAATATGCTTC	3065
Db	2992	AAGGGCTGTTTACCAAGGCTAGGCTGTCTTAAAGCTGATCCCTTTAGTATGTACTTC	3051
QY	3066	CTGTATCAAGGTAAGTATGTGGTGAACAAGGCGAGAA-----CTTCCTTT	3113
Db	3052	CCGGATC--GAGGCACATAAGTGTGCAAAATGAGCGGGGAAATCTTCATTTCTCATTT	3110
QY	3114	TAATTTCTCTCTCTCTTTTATTTTAACAAAATGGTGAAGATGAGGATTAACCTACAAATCA	3173
Db	3111	CTTTCTCTCTCTTAAATATAAATGGCAAAAAAAGATGAAGATTATCTACAAATCA	3170
QY	3174	GACATGGCAAAACAATAATGGCTGTTTGTCTTCATATAACAAGTGCAATTTTTAAAGTGC	3233
Db	3171	GACTTAGCAAAATGATAATGGCTATTTCCGTTCCACATACAAGTGCAATTTTTTAGAGTGC	3230
QY	3234	TGTCTTACTAAGTCTGTTTATTAAGTCTCTCTTTATCTATATGGAATAAAAAAGAGGC	3293
Db	3231	TGTCTTACTAAGTCTGTTTGTGAAGTCTCTCTCTATTTATATGAATAAAGAGAGGC	3290
QY	3294	AGTCATCTTAGCAAAATGACACGTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCTCT	3353
Db	3291	AGTCAATGTTATCAACGGCGTGTCTATTTTCTTAGCTCACCTTGGTCCAACCTGCCTGT	3350
QY	3354	CGATCCCTCTGAGGTATGCCCCATCCAGACTTTTAGGCCATTTCTGATGGAACAGAT	3413
Db	3351	AGAACCTCTCGAGGTATGCCCCCTTCTAAGACTTTTCAGGCCACTCTTGTATGGAATTCGAC	3410
QY	3414	CCC-----TGCCCTGACTGTCCAGCTATCTGAAAGTGGATCAGATTATAAATG	3463
Db	3411	ACCCCTCCCTCAACCCATGACTATCCAGATGCTCTGAATGGGATCAGGTTATAAATG	3470
QY	3464	GATTAATGTAACGTGTTTGGTGTGTG--TCTATCAACCCACAGAGTTCCCTAAACTTG	3522
Db	3471	GATTGCATATGACTGTGTGCTGTGTGTGTGTCAACCTGGACAGAGTTCTCTAAACCTT	3530
QY	3523	CTTCAGTTATAGTAACGTACTGGGTATATTCAATCAGAAAGC-----GCCATAAGTCAGT	3575
Db	3531	CTTTAGTTGTAGCAAGTTCCTGATTCCTCAATCAGAAAGCCCAAGGACATTTGGTGACT	3590
QY	3576	TGAGTATTTGATCCCTAGATAAGAACATGCAAAATCAGCAGAACT--GGTCATACAGGTTA	3634
Db	3591	CGATCAAGGGTTAAACCTTAGGAGAACATGCAAAATAGTAGAACTGGGTGAGACAGGGTA	3650
QY	3635	AGCACCAGGACATTAAGGATTTTATAGATAT-----AATTTAATTTTGTATTGG	3687
Db	3651	AGCACCAGAGATGATAGGATTTTATATATAAATATATAAATTTAATTTTGTATTGG	3710
QY	3688	TTAAGGACAAATTTTGAGAGCAGCAAGCAAAATCTTTT	3723
Db	3711	TTA--TAGACAATTTTGAAGCAAGAGAAATCACT	3744

Db 720 ATATAGCCCTTCGCTCCGGGTCCACTTATGCGCAGACTTATGGCTCGGAATACAC 779
Qy 810 CACGGAGATCATGAACCCCGACTACACCAAGCTGACCATGGACCTTTGGCAGCAGCTAGAT 869
Db 780 CACAGAAATCATGAACCCCGACTATGCAAGCTGACCATGGACCTCGGTAGCAGCGGAT 839
Qy 870 CACGGCTACAGCCACACAGCTCCCTGCGCAGCATCAGTACTCTGTCGGAGGGCTACTCGAG 929
Db 840 CATGGCCACGGCCACAGCTCCCTGCGCAGCTTCAGTACTTCATGGAGGGCTTACCCAG 899
Qy 930 CAATACGAACTCAAGCCCTTCGCTGCTGACCAAAATGCGAG-----CGGCCCTT 977
Db 900 CAGCTGGAACTCAAGCCCTTCGCTGCTGTAACCAATGCGCCCTTCGTCGGGCTCGGCCCTT 959
Qy 978 GATCAAAAGTGGAGGCGGCGGCCAGCTACCATCACTACCATCAACACCAACACCA 1037
Db 960 GATCAAGATGGAAGAGGGTCGGAGCATGGCTACCAACCAACCAACCAACCATCA 1019
Qy 1038 CACACACCAACCATCAACGAGCAGCATCAGAGGCCATCGATTCCTCGAGCCTCCAGGCC 1097
Db 1020 TCATCAACCAACCAACGAG-----CAGCAGCGCTCCATTCCTCCTCCCTCGGCC 1073
Qy 1098 CGAGGACGAGGTGCTGCCAGACCTCCATGTACTTCAAGCAGTCCCAACCGTCCACCC 1157
Db 1074 CGAGGACGAGGTACTGCCAGACCTCCATGTACTTCAAGCAGTCTCCGCGCTAGGCC 1133
Qy 1158 CACCACCGCGCTTCCCGCGCAGCGGGGGCGTTATGGGACGAGGCACTGCGCTCGGC 1217
Db 1134 GACCACTCCAGGCTTCCCGCGCAGCGGGGGCGTGTGGACGAGCTGCGCTTCG 1193
Qy 1218 GCCCGCTGATGCGACCGCGCGCTGCTGAGCCCGCGATGAAGGGCGTCCACGGT 1277
Db 1194 GCTGGCTGATGCTGCTCGGACCGCTGCTGGACCGCGCAGATGAAGGCGATGCGCCCAAT 1253
Qy 1278 GCGCGCGCGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCATCCCGCGCGC 1337
Db 1254 GCGCGCTGCTGCGGCTTCCGATCTTCAAGCCCTCACCGCCACACCTCCCGCGC 1313
Qy 1338 CAGCCCGCGCGCGCCACCATCGGCTACGACCGCGCGCTGCGCGCTCAGCCT 1397
Db 1314 CAGCCAGCGCGCGCCACCATCGGCTATGACCCCAACCGCGCAGCTGCGCTCAGTCT 1373
Qy 1398 GCGCTGGAGCGCAGCGCGCGCGCAGCCAGCGCGCGCTTGAGAGCACCCTGA 1457
Db 1374 ACCCTGGAGCGCGCGCGCGCGCGCGCAGCTGCTGCGCTCGAGGGCCATCCGTA 1433
Qy 1458 CCGGCTGCGCTGGCCAAAGAGGCGCGCGCGCTTCCCGCTCGCGCTCAGCGC 1517
Db 1434 CCGGCTCCCGCTGGCCAAAGAGGCGCGCGCGCTTCCCGCTCGCGCTCAGCGC 1493
Qy 1518 CTCCTTACCGCTCCAGCTTGTGGCGGAGAGTCCAGCTGCGCTGCGCGCCAGCAG 1577
Db 1494 GTCCCTTACCGCTCCAGCTTGTGGAGAGAGCCTCCAGCTTACCATCGCCACCAATAG 1553
Qy 1578 GAGCTGCTGTGGGCGAGGCGACGTGTGCGGTGGGGGCAACCGCGCTTCGCGCA 1637
Db 1554 GAGCTCATATCCGCGGAGGCGACGTGTGTGTGGGGGCAAAATGCTGCTGCGCGCA 1613
Qy 1638 CTACGGCGTGGAACTGCGAGGCTGCAAGGCTTTTCAAGAGAACAGTGCAGAAAA 1697
Db 1614 CTACGGAGTCCGACCTGCGAGGCTGCAAGGCTTCTTCAAGAGAACGCTGCAGAAAA 1673
Qy 1698 TGCAAAATATGTTGCTGGCAATAAAACTGCCCAGTAGACAAGAGAGCTCGAAACCG 1757
Db 1674 CGCAAAATATGTTGCTGGCAATAAAACTGCCCGGTAGACAAGAGAGCTCGAAATCG 1733
Qy 1758 ATGTAGTACTGTGCAATTTCAAGAGTGTCTCAGTGTGGAATGGTAAAAAGAGTGTCCG 1817
Db 1734 ATGTAGTACTGTGCAAGTGTCTCAGTGTGCGGATGGTGAAGAGAGTGTGTGCG 1793
Qy 1818 TACAGATGCTGAAGAGGAGAGGCTGCTGCTCCCAACCAAGAGCCATTACA 1877

Db 1794 TACAGATGCTGAAAGGGAGGAGGCTGCTGCTCCAAACCAAGAGCCCACTACA 1853
Qy 1878 ACAGGAACCTTCTCAGCCCTCTCCACTTCTCTCTCAATCTGATGATGAATGCCCTTGT 1937
Db 1854 ACAGGAGCCCTCGCAGCCCTCCACCATCTCTCGATCTGTATGATGAACCCCTTGT 1913
Qy 1938 CCGAGCTTTAAACAGACTCAACCCAGAGATCTTGATTTATCCAGATCTGTGCCACTGA 1997
Db 1914 CCGAGCTTTAAACAGAGCGAACGCCAGAGACCTTGATTTACTCCAGATCTGTCCCAACCGA 1973
Qy 1998 CCGAGCTGTGCGAGCAGACAGATGCTGAGCATGTGCAACAAATTTCAACCTCTCTGACGC 2057
Db 1974 CCGAGCCACTGCGGGCAGACAGCGTGAGCAGCTGCGAGCAGTCTTCAACCTCTGAGCGC 2033
Qy 2058 CTCCATTTGATGATTCAGAAAGCTGGGCGAGAAAGATTCGGGATTTACTGATCTCCCAA 2117
Db 2034 CTCCATCGAGCTGTCCAGAAAGCTGGGCGAGAAAGATTCCTCGGATTCATGATCTCCCAA 2093
Qy 2118 AGAAGATCAGACATTTATTTGATCAGCCCTTTTGGAGCTGTTTGTCTCAGACTTC 2177
Db 2094 AGAAGATCAGAGCTTACTTTATAGAATCAGCCCTTTTGGAGCTGTTGTTCTTAGACTTC 2153
Qy 2178 CATCAGGTCAAACTGCTGAAAGATAAGTTTGTGTTCTGCAATGGACTTGTCTCGATCG 2237
Db 2154 TATCAGGTCAAACTGCTGAAAGATAAGTTTGTGTTCTGCAATGGACTTGTCTCGACCG 2213
Qy 2238 ACTTCAGTCCCTGCTGGGATTTGGGAGTGGTCTGATCTATTTAAAGACTTTTCTTTAAA 2297
Db 2214 ACTTCAGTCCCTGCGGATTTGGGAGTGGTCTGACTCCATTTAAAGACTTTTCTTTAAA 2273
Qy 2298 TTTGCGAGCTGAACTTGTATCCAGCTTAGCTTCCCTGCTGAGCAGTGAAGATGAT 2357
Db 2274 TTTGCGAGCTGAACTTGTATCCAGCTTAGCTTCCCTGCTGAGCAGTGAAGATGAT 2333
Qy 2358 CACAGAAAGACATGGGTTTAAAGAACCAAGAGAGTTCGAAGAGCTATGCAACAAAGATCAC 2417
Db 2334 CACAGAGCAGATGGGTTTAAAGAACCAAGAGAGTTCGAAGAGCTATGCAACAAAGATCAC 2393
Qy 2418 AAGAGTTTTAAAGACCAACAGATGAAGGACAGGCTCTGGAGCCCAACAGTCCAGGT 2477
Db 2394 AAGCAGCTTAAAGGACCAACAGAGGAGGAGGAGGCTCTGGAGCCCTCAGAGCCCAAGGT 2453
Qy 2478 CTTGGGTGCTGTTAGACTGAGGAAGATCTGCAACCTTCATTTGCAAGCTCTTCTCGGACAC 2537
Db 2454 CTTTCGCGCCTGTTGGAACTGAGGAAGATCTGCAACCGAGGCTCCAGGCTATCTCTA 2513
Qy 2538 CTTGAAGCTGGAAGACTTGGTGTCTCCACTTCCATTTGCAAGCTCTTCTCGGACAC 2597
Db 2514 CTTGAAGCTGGAAGACTTGGTGTCTCCACTTCCCACTTCTGTCATCGACAAGCTCTTCTCTGATAC 2573
Qy 2598 CTTA CTTTCTAATCAGGAGCAG - TGGAGAGTGAAGTCTGCTCTCTCTAGCACTGCT 2656
Db 2574 CTTGCTTCTGAGCAGGGGAGCTGAGCAGAGAGCTTCTGCTCTGCTGCGACCTGCTC 2633
Qy 2657 TGTACGCGCAAGGGATAGTTTGGAAACCTATCATTTCTGCTCTCTTCTTAAGGGA 2716
Db 2634 ATTAAGTGAAGGAAAGATGGGTTTGAACACCT - GCCCTCTATCTCTCTCAGGGGAA 2692
Qy 2717 AAGCAGCTCTGTAGAAAGCAAGACTTCTTTTTTTTCTGCTCTCTTCTTCTTCAACAC 2776
Db 2693 AAGCAGCTCTTCTGAGAAAGCAAGACTT - TTTTTTCTGCGACCTTCTTCTTCAACAC 2751
Qy 2777 TAAAGCCAGAAACTTGCAGAGTATTTGTTGGGTTGTTTATATTTATTTAGGATGGG 2836
Db 2752 TAAAGCCAGAAACCTTGCAGAGTATTTGTTGGGTTGTTTATATTTAGGCTTTGGT 2811
Qy 2837 GATGGGTTGGAGGGGT - TATAGTTTATGAGGTTTTTCTTAAGAAATGCTTAAACAG 2894
Db 2812 GGTGGGCTGGAGGGGTAAATAGTTTATGAGGCTTTTCTAAAGAAATTTGCTGAGGAAG 2871
Qy 2895 CACTTTTGGACAAATGCTATCCAGCAG - - - - - GAAAAAAGGATAATATACTGTTTT 2948
Db 2872 CACTTTTGGATGATGCTATCCAGCAGTGGGTTGGGAGGAAGGATATATACTGTTTT 2931

3114	TAATTTCCTTCCTTCCTTTATTTTAAACAAATGGTGTGAAGATGGAGGATTACCTACAAATCA	3173
3111	CTTTCTCTCTTTAAAAATAAATGGCAAAAAAAGATGGAAGATTATCTACAAATCA	3170
3174	GACATGGGCAAAACAATAATGGCGTGTTCCTCCATAAACAAGTGCAGTAATTTTTTAAAGTGC	3233
3171	GACTTAGCAAAATGATAATGGCTATTTCGCTTCACATACAGTGCATTTTTTTTAGNGTGC	3230
3234	TGTCCTTACTAAGTCTTTGTTTTATTAACCTCCTTTATTCATATGGAAAAATAAAGAGAGC	3293
3231	TGTCCTTACTAAGTCTTTGTTTGTGAACCTCCCTCAATTTATATGAAAAATAAGAAGAGAGC	3290
3294	AGTCATGTTTAGCAAAATGACACGTTATATATCCCTAGCAGAGGCTGTGTTCACTTCCCTCT	3353
3291	AGTCATGTTATCAACGGCGTGTCTAATTTCTTAGCTCACCTTGGTCCACCTGCCCTCT	3350
3354	CGATCCCTCTCTGAGGTATGGCCCATCCAAGACTTTTATAGGCCATTTCTTGATGGAAACAGAT	3413
3351	AGAACCTCTCGGAGGTATGGCCCTTCTAAGACTTTTCAGGCCACTCTTGATGGAATTCGCAC	3410
3414	CCC-----TGCGCTGACTGTCAGCTATCTCTGAAAGTGGATCAGATTATAAATCG	3463
3411	ACCCCTCCCTCAACCCATGACTATCCAGATGTCCTGAAATGGGGATCAGGTTATAAATG	3470
3464	GATTACATGTAACTGTTTTGGTTGTGT-TCTATCAACCCCAACAGAGTTTCCCTAAACTTG	3522
3471	GATTGCATATGACTGTGTTCGCTGTGTGTGTGTCAACCTGGACAGAGTTCTCTAAACCTT	3530
3523	CTTCAGTTATAGTAACTGACTGTGTATTTTCATTTCAGAACG-----GCCATAAGTCAGT	3575
3531	CTTTAGTTGTGACAAAGTTCTCTGATTCCTCATTCAGAAAGCCCAAGAGCATTTGGTGACT	3590
3576	TGAGTATTTTCATCCCTAGATAAAGAACATGCAAAATCAGCAGGAACT-GGTCATACAGGGTA	3634
3591	CGATCAAGGGTTAAACCCCTAGGAGAAACATGCAAAATAAGTAGGAACTGGGTTCAGACAGGGTA	3650
3635	AGACACNGGACAAATAGGATTTTTTATAGATAT-----AATTTAATTTTTTGTTATTGG	3687
3651	AGACCCAGAGATGATTAAGGATTTTATATATAAATAATATAATAAATAATTTTTGTTATTGG	3710
3688	TTAAGGAGACAAATTTTGGAGAGCAAGCAAAATCTTTT	3723
3711	TTA--TAGACAAATTTTGCAAAGCAAGAGATTCATCT	3744

```

RESULT 14
US-10-191-803-209
; Sequence 209, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191.803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 209
; LENGTH: 4400

```

[illegible]

Qy	978	GATCAAAGTGGAGGAGGGGGCGCCACGCTACCATTCAACAACCAACAACCAACCA	1037
Db	960	GATCAAGATGGAAGAGGGTGCGCAGCATGGCTACCAACACCAACCAACCAACCATCA	1019
Qy	1038	CCACCAACCATCACCAAGCAGCAGCATCAGCAGCCATTCATTCTCTCCAGCCTCCAGCCC	1097
Db	1020	TCAATCAACAACAACACAGAG-----CAGCAGCGTCCATTCTCTCTCTGTGCCCC	1073
Qy	1098	GGAGGACGAGGTGCTGCCACAGCACTCCATTGTACTTCAAGCAGTCCCAACCGTCCAACCCC	1157
Db	1074	CGAGGACGAGGTACTGCGCCAGCACTCCATGTACTTCAAGCAGTCTCCGCGCTTAGCGC	1133
Qy	1158	CACCAACGCGGCCTTTCCCCCGCAGCGGGGGGGTATGGGACGAGGCATCTGCCTCGCC	1217
Db	1134	GACCACTCCAGGCTTCCCCCGCAGCGGGGGGGCTGTGGGACGACGAGTGCCTCTGCTG	1193
Qy	1218	GCCCGGCTGCATCGCACACCGCGCGCTGTGGACCGCCCGCATGAAGGCGGTGCCACGGT	1277
Db	1194	GCTTGGCTGCATGCCTTCGGGACCGCTGTGGACCGCCGACATGAGAGCAGTGCCTCCAT	1253
Qy	1278	GGCGCGCGCGCTTCCCGCTCTTCCACTTCAAGCCCTTCGCGCGCATCCCCCGCGCC	1337
Db	1254	GGCGCTGCTGCGCGCTTCCCGATCTTCTTCAAGCCCTCACCGGCAACCGTCCCGGCCC	1313
Qy	1338	CAGCCCGCGCGGGCCACCACTCGGCTACGACCCGACGGCGGCTGCGCGGCTCAGCCT	1397
Db	1314	CAGCCCAAGCGGGGCCCCACCACTTGGCTATGACCCCAACGGCGGCACTGCTCAGTCT	1373
Qy	1398	GCCGCTGGAGCCGACGACGCGCGCGGCGCAGCAGGCGCGCGCTTGGAGGCCAACCGTA	1457
Db	1374	ACCCCTGGAGCGCGCGCGCGGCGAGCGCAAGCTGTCTGCGCTCGAGGGCCATCCGTA	1433
Qy	1458	CGGCTGCCCTGGCCAAAGAGGGCGGCCCGCTGGCTTCCGCGCTCTCGGCTCAAGCC	1517
Db	1434	CGGCTCCCGCTGGCCAAAGAGGACGGGCCAAGTTGACCTTCCCTCGGCTGGGCTCAAGC	1493
Qy	1518	CTCCCTTACCGGCTCAGCCTGTGTGGGGAGAGTCCAGCCTGCGGTCGCGGCGCCAGCAG	1577
Db	1494	GTCCCCCTACCGCGCTCAGCCTGTGTGGGAGAGAGCCCGAGCTTACCATCGGCAACCCAA	1553
Qy	1578	GAGCTCGTCTGTGGCGAGGCGACGTGTGCGGTGTGGGGGACAAAACCGCGCTGCCAGCA	1637
Db	1554	GAGCTCATCATCCGCGAGGGCAGTGTGTGTGTGGGGACNAAGCTGCTGCCAGCA	1613
Qy	1638	CTACGGCGTGGGAACCTTCGAGGGCTGCAGAGGGCTTTTTCAAGAGAACAGTCGAGAAAA	1697
Db	1614	CTACGGAGTCCGCACTTCGAGGGCTGCAGAGGGCTTCTTCAAGAGAACGGTCGAGAAAA	1673
Qy	1698	TGCAAAATATGTTTTGCTGTGCAATAAAACTGCCCAGTAGACAAGAGAGCTCGAAACCG	1757
Db	1674	CGCAAAATATGTTTTGCTTGGCAATAAAACTTGCCCGGTAGACAAGAGAGCTCGAATCG	1733
Qy	1758	ATGTCAGTACTGTGCATTTCAGAAGTGTCTCAGTGTGTGAATGGTAAAAAGATGTGTCG	1817
Db	1734	ATGTCAGTACTGCAAGTTTTCAGAAGTGTCTCAGTGTGCGGATGGTGAAGAGATGTGCG	1793
Qy	1818	TACAGATAGTCTGAAAGGGAGGAGAGTGTGTGCTCTCCAAACCAAGAGGCCCATACCA	1877
Db	1794	TACAGATAGTCTGAAAGGGAGGAGAGTGTGTGCTCTCCAAACCAAGAGGCCCATACCA	1853
Qy	1878	ACAGGAACCTTCTCAGCGCTCTCCACCTTCTCCCTCAATCTGCAATGATGAATGCCCTGT	1937
Db	1854	ACAGGACCTCTCGCAGCGCTTCCCAACCATCTCTCTCCGATCTGTATGATGAACGCCCTGT	1913
Qy	1938	CCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTTATTTCCAGATACTGTCCCACTGA	1997
Db	1914	CCGAGCTTTAAACAGACGCAACGCCCCAGAGACTTGAATTAATCCAGATACTGTCCCAACCA	1973
Qy	1998	CCAGGCTGTGCGGCGACAGATGCTGAGCATGTGCAACAATTTCTAACTCTCTGACGACG	2057
Db	1974	CCAGGCACTGTGCGGCGACAGACGCTGAGCAGCTGTGAGCAGTTCATCAACTCTTGTGAGCG	2033

Qy	2058	CTCCATTGATGTTATCCAGAAAGCTGGGCAGAAAAGATTCCGGGATTTTACTGATCTCCCCAA	2111
Db	2034	CTCCATCGAGTGTCCAGAAAGCTGGGCAGAAAAGATCCCGGATTCACGTGATCTCCCAA	2093
Qy	2118	AGAAGATCAGACATTAATTATCAATCAGCCCTTTTGGAGCTGTGTGTCTCTCAGACTTTC	2177
Db	2094	AGAAGATCAGACGTTACTTATAGAAATCAGCCCTTTTGGAGCTGTGTGTCTCTAGACTTTC	2153
Qy	2178	CATCAGGTCAAAACACTGCTGSAAGATAAGTTTGTCTTGCATGGACTTGTCTCTGCATCG	2237
Db	2154	TATCAGGTCAAAACACTGCTGSAAGATAAGTTTGTGTCTTGCATGGACTTGTCTCTGCACCG	2213
Qy	2238	ACTTCAGTGCCTTCGTGGATTTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTTAAA	2297
Db	2214	ACTTCAGTGCCTTCGTGGATTTTGGGAGTGGCTCGACTCCAATTAAAGACTTTTCTTTAAA	2273
Qy	2298	TTTTCAGAGCTCGAACTCTGATATCAAGCCTTTAGCTGCTGTCAGCACTGTCAGCACTGAGCATGAT	2357
Db	2274	TTTTCAGAGCTCGAACTCTGATATCAAGCCTTTAGCTGCTGTCAGCACTGAGCATGATGAT	2333
Qy	2358	CACAGAAGACATCTGGGTTAAAGAACCAGAGAGATCGAAGACTATGCAACCAAGATCAC	2417
Db	2334	CACAGAGCAGATGGGTTAAAGAACCAGAGAGATGGAGAGCTATGCAACCAAGATCAC	2393
Qy	2418	AAGCAGTTTAAAGACCAACAGATGAAAGGACAGGCTCTGGAGCCACCGAGTCCAAGGT	2477
Db	2394	AAGCAGTTTAAAGACCAACAGAGGAGGACAGGCTCTGGAGCCCTCAGAGGCCAAGGT	2453
Qy	2478	CCTGGGTGCCCTGTTAGAACTGAGGAGAGATCTGCACCCTGGGGCTCCAGCGCATCTTCTA	2537
Db	2454	CCTTGGGCACTGGTGAATCTGAGGAAGATCTGCACCCAGGGCTCCAGGGTATCTTCTA	2513
Qy	2538	CCTGAAGCTCGAAGACTTGGTGTCTCCAACCTTCATCATTTGCAAGCTCTTCTCTGGAAC	2597
Db	2514	CCTGAAGCTCGAAGACTTGGTGTCTCCAACCTTCTGTCGACGAAGCTCTTCTCTGATAC	2573
Qy	2598	CCTACCTTTCTAATCAGAGCAG-TGAGCAGTGAGTGCTGCTCCTCTCTAGCACTGCT	2656
Db	2574	CCTGCCTTTCTGAGCAGGGGAAGCCTGAGCAGAGAGCTACTTGTCTCTGCTGGCACTGGTC	2633
Qy	2657	TGCTACGCAAGAGGATAGTTTGGAAACCTATCATTTCTGTCTCTTCTTAAAGGA	2716
Db	2634	ATTAAGTGACAAAGGATGGGTTTGAACACT-GCCCCCTTATCTCTCTCCAGGGGA	2692
Qy	2717	AAAGCAGCTCCTGTAGAAACAAAGACTTTCTTTTCTTTTCTGTGCTCTTTTCTTACAAC	2776
Db	2693	AAAGCAGCTCCATAGAAAGCAAGACTTT-TTTTTTCTCTGGCACTTCTTCTACAAC	2751
Qy	2777	TAAAGCCAGAAAACCTTCAGAGTATTGTGTGGGGTGTGTTTATATTATTAGGATTTGGG	2836
Db	2752	TAAAGCCAGAAAACCTTCAGAGTATTGTGTGGGGTGTGTTTATATTATTATTAGGCTTTGGT	2811
Qy	2837	GGAATGGGTGGGAGGGGT--TATAGTTCAATGAGGTTTCTAAGAAATGCTTAAACAAG	2894
Db	2812	GGGTGGGCTGGGAGGGGTAAATAGTTTCAATGAGGCTTTTCTAAGAAATGCTGACGAAG	2871
Qy	2895	CACCTTTTGGACAATGCTATCCACGACGAG-----GAAAAAAGGATAATAACTGTGTTT	2948
Db	2872	CACCTTTTGGATGATGCTATCCAGCAGTGGGTGGGAGAAAGGATAATATACTGTGTTT	2931
Qy	2949	AAAA--CTCTTTCTGGGAAATCCAAATTATAGTTGCTTTGTATTATTAATAAACAAGACAGCC	3006
Db	2932	AAAAACTCTTTTCCGGGGAAATAGCATATGCTTGTGCTTTGTATTATTAATAAACAAGACAGCC	2991
Qy	3007	AAGGTTG--TTCGCCAGGTTAGATGTGCTTAAAGATTGGTCCCTTGAATAATATGCTTC	3065
Db	2992	AAGGGCTGTTTATCCAGGGTAGGGCTGTGTCTTAAAGCTGATCCCTTTAGTATGTACTTC	3051
Qy	3066	CTGTATCAAAAGGTACGTATGTGTGCAAAACAAGGCAGAAA-----CTTCTCTTT	3113
Db	3052	CCGATC-GAGGACATAGTGGTGCNAATGAGCGGGGGAATTCCTTCATTTCTTCATTT	3110
Qy	3114	TAATTTCTTCTTCTCTTTATTTTAAACAAATGGTGAAGATGAGGAGGATTAACCTACAAATCA	3173

Tue Mar 22 09:54:22 2005

us-10-608-863-1.rnpb

Page 30

Search completed: March 22, 2005, 01:21:50
Job time : 2024 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2005, 08:40:51 ; Search time 11544 Seconds
(without alignments)
12510.034 Million cell updates/sec

Title: US-10-608-863-1
Perfect score: 3794
Sequence: 1 ataaatgcgcgcgcagagaga.....gtgtgactactaggagagat 3794

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	879.8	23.2	996	1	AJ814207
2	796	21.0	1087	4	BM557376
3	769.4	20.3	1032	4	BM549822
4	752.4	19.8	819	5	BO431513
5	696.4	18.4	852	7	CK778210
6	627.8	16.5	671	6	CB445232
7	617	16.3	757	6	CD521819
8	597.2	16.1	679	6	CB465379
9	597.2	15.7	635	6	CB447412
10	593	15.4	940	5	BU436288
11	563	14.8	635	6	CD638747
12	561	14.8	794	6	CD522275
13	556.2	14.7	689	6	CB447773
14	555.4	14.6	802	7	CK635592
15	552.6	14.6	644	6	CB445564
16	537	14.2	735	6	CD523420
17	524.4	13.8	713	7	CF536097
18	520.4	13.7	586	5	BX925243
19	518.2	13.7	708	6	CD349619
20	495.4	13.6	533	2	BF086936
21	495.4	13.1	1797	9	AY417753
22	490.8	12.9	1797	9	AY417755
23	488.8	12.9	588	2	AW491078
24	480.8	12.7	540	5	BX920563

25	474.8	12.5	574	1	AI145834
26	474.4	12.5	2759	3	AK087495
27	461.6	12.2	493	2	BF604409
28	429.2	11.7	478	7	CR548239
29	429.4	11.3	537	7	N40404
30	428.8	11.3	441	1	AI911092
31	426.4	11.2	433	4	BG235965
32	425.6	11.2	1699	9	AY417754
33	422.8	11.1	525	2	BB758614
34	414.4	10.9	564	2	BF019466
35	411	10.8	500	2	AW490389
36	390	10.3	390	1	AV645408
37	390	10.3	390	1	AV645902
38	385.8	10.2	452	7	CR555348
39	376.4	9.9	483	6	CB441272
40	373.2	9.8	487	4	BI399884
41	372.6	9.8	451	1	AI327473
42	372.2	9.8	577	2	BE648145
43	365.8	9.6	407	5	BO326341
44	363.8	9.6	2435	3	CR590328
45	351.2	9.3	428	1	AI323943

ALIGNMENTS

RESULT 1
AJ814207
LOCUS AJ814207 KN206 Bos sp. cDNA clone C0005204p19, mRNA linear EST 13-SEP-2004
DEFINITION AJ814207
ACCESSION AJ814207
VERSION AJ814207.1 GI:51881683
KEYWORDS EST.
SOURCE Bos sp.
ORGANISM Bos sp.

REFERENCE 1 (bases 1 to 996)
AUTHORS McGuire,K. and Glass,E.J.
TITLE Unpublished ESTs, McGuire and Glass
JOURNAL Unpublished (2004)
COMMENT Contact: McGuire K
Genomics and Genetics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
Single pass sequencing. Bases called and trimmed with phred v0.020425 c. Vector identified by cross match with the -minscore 20 and -minmatch 12 options. Vector: BluescriptII(SK+) R. Site 1: EcorV(lost) R. Site 2: NotI Seq Primer: T7 Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Sahiwal) cattle subjected to various stimuli, including infection with the protozoan parasite Theileria annulata.

FEATURES

location/Qualifiers
1..996
/organism="Bos sp."
/mol_type="mRNA"
/db_xref="taxon:29061"
/clone="C0005204p19"
/tissue_type="blood"
/cell_type="bovine monocyte"
/clone_lib="KN206"
/note="Vector: pBluescriptII(SK+); Site 1: EcorV(lost); Site 2: NotI; Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Sahiwal) cattle subjected to various stimuli, including infection with the protozoan parasite Theileria annulata"

ORIGIN

Query Match 23.2%; Score 879.8; DB 1; Length 996;
Best Local Similarity 94.9%; Pred. No. 1.3e-202;
Matches 919; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

1691 AGAAAAATGCAAAATATGTTTGGCTGGCAAAATAAAATGCCCCAGTAGACAAGACGTC 1750
Db |||||
25 AGAAAAATGCAAAATATGTTTGGCTGGCAAAATAAAATGCCCCAGTAGACAAGACGTC 84
QY |||||
1751 GAAACCGATGTCAGTACTGTCGATTTGCAAGTGTCTAGTGTTCGATGGTAAAGAG 1810
Db |||||
85 GAAACCGATGTCAGTACTGTCGATTTGCAAGTGTCTAGTGTTCGATGGTAAAGAG 144
QY |||||
1811 TTGTCGCTACAGATAGTCTGAAGGGAGAGAGTGTCTGCTGCTTCCAAACCAAGAGGCC 1870
Db |||||
145 TTGTCGCTACAGATAGTCTGAAGGGAGAGAGTGTCTGCTGCTTCCAAACCAAGAGGCC 204
QY |||||
1871 CATTACAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAATTCGATGATGAATG 1930
Db |||||
205 CATTACAGCAGGAAGCTCTCAGCCCTCTCCACCTTCTCCTCCGATCTGTATGATGAATG 264
QY |||||
1931 CCCTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTGATTTTCCAGATAGTCTGTC 1990
Db |||||
265 CCCTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTGATTTTCCAGATAGTCTGTC 324
QY |||||
1991 CCACTGACCAGGCTGCTGAGGACAGATGCTGAGCATGTGCAAAATTTCTACAACCTCC 2050
Db |||||
325. CCACTGACCAGGCTGCTGAGGACAGATGCTGAGCATGTGCAAAATTTCTACAACCTTC 384
QY |||||
2051 TGACAGCCTCCATTGATGATCCAGAGCTGGGCGAAGAGATTCGGGATTTACTGATC 2110
Db |||||
385 TAACAGCCTCCATTGATGATCCAGAGCTGGGCGAAGAGATTCGGGATTTACTGATC 444
QY |||||
2111 TCCCAAGAGATCGACATTTACTTATTGATTCAGCTTTTGGAGCTTTTGGCTTCTCA 2170
Db |||||
445 ACCCCAAAGAGATCAGACATTTACTTATAGATTCAGCTTTTGGAGCTTTTGTCTCA 504
QY |||||
2171 GACTTTCCATCAGTCAACACTGCTGAAGATAAGTTTGTCTTGCATGACCTGTGCC 2230
Db |||||
505 GACTTTCCATCAGTCAACACTGCTGAAGATAAGTTTGTCTTGCATGACCTGTGCC 564
QY |||||
2231 TGACATGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTGACTCTATTAAAGACTTTT 2290
Db |||||
565 TGACATGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTGACTCCATTAAAGACTTTT 624
QY |||||
2291 CCTTAAATTTGAGAGCTTGAACCTTGATATCCAGCCTTAGCCTGCTGCTGAGCACTGA 2350
Db |||||
625 CCTTAAATTTGAGAGCTTGAACCTTGATATCCAGCCTTAGCCTGCTGCTGAGCACTGA 684
QY |||||
2351 GCATGATACACAGAAACATAGGTTTAAAGAACCCAAAGAGAGTGCAGAGACTATGCAACA 2410
Db |||||
685 GCATGATACACAGAAACATAGGTTTAAAGAACCCAAAGAGAGTGCAGAGACTATGCAACA 744
QY |||||
2411 AGATCAACAGCAGTTTAAAGAACCCACAGAGTAAGGGAAGGCTGTGGAGCCCACTGAGT 2470
Db |||||
745 AGATCAACAGCAGTTTAAAGAACCCACAGAGTAAGGGAAGGCTGTGGAGCCCACTGAGT 804
QY |||||
2471 CCAAGTCTGGTGGCTGCTGAGTACAGTACAGAGATCTGCACCTGGGCTCCAGCGCA 2530
Db |||||
805 CCAAGTCTGGTGGCTGCTGAGTACAGTACAGAGATCTGCACCTGGGCTCCAGCGCA 864
QY |||||
2531 TCTTCTACCTGAGCTGGAGAGCTTGGTGTCTCCACTTCCATTCATTGACAAGCTCTTCC 2590
Db |||||
865 TCTTCTACCTGAGCTGGAGAGCTTGGTGTCTCCCACTTCCATTCAGCAAGCTCTTCC 924
QY |||||
2591 TGGACACCTTACTTTTCTTAATCAGGAGAG-TGGAGCAGTGAAGCTGCTTCTCTCTAGC 2649
Db |||||
925 TGGACACCTTACTTTCTGAGCAGGAGAGCTGAGCAGGAGAGCTGCTCATCTGCTAGC 984
QY |||||
2650 ACTGCTT 2657
Db |||||
985 AGCTGCTT 992

RESULT 2
BM557376 1087 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6561893 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5547838
DEFINITION

5', mRNA sequence.
BM557376
VERSION BM557376.1 GI:18799320
EST.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (Bases 1 to 1087)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM12255 row: m column: 23
High quality sequence stop: 624.
Location/Qualifiers
1..1087
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5547838"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 21.0%; Score 796; DB 4; Length 1087;
Best Local Similarity 96.7%; Pred. No. 3.3e-182;
Matches 878; Conservative 0; Mismatches 20; Indels 10; Gaps 6;
QY 2869 GGTTCCTTAAGAAATGCTTAACAAAGCACTTTGGACAATGCTATCCAGCAGGAGAAAAA 2928
Db |||||
1 GGTTCCTTAAGAAATGCTTAACAAAGCACTTTGGACAATGCTATCCAGCAGGAGAAAAA 60
QY 2929 AAGGATAATAAATGCTTTTAAACTCTTCTGGGGAATCCAATATATAGTTGCTTTGTAT 2988
Db |||||
61 AAGGATAATAAATGCTTTTAAACTCTTCTGGGGAATCCAATATATAGTTGCTTTGTAT 120
QY 2989 TTAATAAACAAGAACAGCCAGGGTTGTTCCGAGGGTAGGATGTCTTTAAAGATTGGTC 3048
Db |||||
121 TTAATAAACAAGAACAGCCAGGGTTGTTCCGAGGGTAGGATGTCTTTAAAGATTGGTC 180
QY 3049 CTTTGAATAATGCTTCTCTGATCAAGGTAGTATGTGTCGCAACAGGAGCAAACTT 3108
Db |||||
181 CTTTGAATAATGCTTCTCTGATCAAGGTAGTATGTGTCGCAACAGGAGCAAACTT 240
QY 3109 CTTTGAATAATGCTTCTCTGATCAAGGTAGTATGTGTCGCAACAGGAGCAAACTT 3168
Db |||||
241 CTTTGAATAATGCTTCTCTGATCAAGGTAGTATGTGTCGCAACAGGAGCAAACTT 3300
QY 3169 AATCAGCATGGCAAAACAATGCTGTTGCTTCCATAAACAAGTGCATTTTAA 3228
Db |||||
301 AATCAGCATGGCAAAACAATGCTGTTGCTTCCATAAACAAGTGCATTTTAA 360
QY 3229 AGTGTGCTTACTAAGTCTTGTATTAACTCTCTTATCTATATGGAATAAAG 3288
Db |||||
361 AGTGTGCTTACTAAGTCTTGTATTAACTCTCTTATCTATATGGAATAAAG 420
QY 3289 GAGGAGCATGTTAGCAAAATGACACGTTAATATCCCTAGCAGAGGCTGTGTTCACTTC 3348
Db |||||
421 GAGGAGCATGTTAGCAAAATGACACGTTAATATCCCTAGCAGAGGCTGTGTTCACTTC 480

QY 3349 CCTGTCGATCCCTTCTGAGGTATGGCCATCCAGACTTTTAGGCCATTTCTGATGGAAC 3408
|||||
Db 481 CCTGTCGATCCCTTCTGAGGTATGGCCATCCAGACTTTTAGGCCATTTCTGATGGAAC 540
|||||
QY 3409 CAGATCCCTGCTGACTGTCCAGCTATCTGAAAGTGGATCAGATTATATAAACTGGATTA 3468
|||||
Db 541 CAGATCCCTGCTGACTGTCCAGCTATCTGAAAGTGGATCAGATTATATAAACTGGATTA 600
|||||
QY 3469 CATGTAACCTGTTTGGTGTGTTCTATCAACCCACAGATTCCTTAAACTTGTCTAG 3528
|||||
Db 601 CATGTAACCTGTTTGGTGTGTTCTATCAACCCACAGATTCCTTAAACTTGTCTAG 660
|||||
QY 3529 TTATAGTAACCTGCTGATATCTTCAAGAGCGCCATAAGTCAG-TTGAGTATTGAT 3587
|||||
Db 661 TTATAGTAACCTGCTGATATCTTCAAGAGCGCCATAAGTCAG-TTGAGTATTGAT 720
|||||
QY 3588 CCTAGATAAGAACATGCAAAATCAGCAGGAATCTGGTCATACAGGGTAAG-CACCAGGGAC 3646
|||||
Db 721 CCTAGATAAGAACATGCAAAATCAGCAGGAATCTGGTCATACAGGGTAAG-CACCAGGGAC 780
|||||
QY 3647 AATAAGGATTTTATAGATATAATTTAA--TTTTGTTATTGTTAAGGAG-ACAATTTT 3703
|||||
Db 781 AATAAGGATTTTATAGATATAATTTAAATTTTTCCTATTGGCTTAAGGAGAACATTTT 840
|||||
QY 3704 GGAGAGCCAGCAAA-----TCTTTTAAATAATAGTATGAAT-CTGTAATCTAGAAAGAT 3758
|||||
Db 841 GGAGAGCCAGCAAAATCTCTTTTAAATAATAGTATGAATGGGGAATACCAAAAAAC 900
|||||
QY 3759 TTAATAAA 3766
|||||
Db 901 TTTAAGAA 908
|||||

RESULT 3

BM549822
LOCUS BM549822 1032 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6544050 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745831
5', mRNA sequence.

ACCESSION BM549822

VERSION 1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1032)

NIH-MGC <http://mgi.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM12769 row: g column: 16

High quality sequence stop: 651.

Location/Qualifiers

1..1032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5745831"

/tissue_type="leukocyte"

/lab_host="DH10B"

/clone_lib="NIH MGC 118"

/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed

and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 20.3%; Score 769.4; DB 4; Length 1032;
Best Local Similarity 94.4%; Pred. No. 9.7e-176;
Matches 820; Conservative 0; Mismatches 46; Indels 3; Gaps 2;

QY 2360 CAGAAGACATGGTTAAAGAACCAAGAGAGTGAAGAGCTATCAACAGATCACA 2419
Db 5 CCGGAATTCCTCGGATATAAGAACCAAGAGAGTGAAGAGCTATCAACAGATCACA 64
|||||
QY 2420 GCAGTTTAAAGACCAACAGAGTAAGGGAAGGCTCTGGAGCCACCGAGTCCAAGGTCC 2479
Db 65 GCAGTTTAAAGACCAACAGAGTAAGGGAAGGCTCTGGAGCCACCGAGTCCAAGGTCC 124
|||||
QY 2480 TGGGTGCCCTGTAGAACTGAGAGAGATCTGCACCTCTGGGCTCTCAGCGCATCTTCTACC 2539
Db 125 TGGGTGCCCTGTAGAACTGAGAGAGATCTGCACCTCTGGGCTCTCAGCGCATCTTCTACC 184
|||||
QY 2540 TGAAGCTGGAAGACACTTGGTGTCTCCACCTTCCATCATTTGACAGCTCTTCTGGACACCC 2599
Db 185 TGAAGCTGGAAGACACTTGGTGTCTCCACCTTCCATCATTTGACAGCTCTTCTGGACACCC 244
|||||
QY 2600 TACCTTTCTAATCAGGAGCAGTGGAGCAGTGGAGTGGCTCTCTCTAGCACCCTGTGTC 2659
Db 245 TACCTTTCTAATCAGGAGCAGTGGAGCAGTGGAGTGGCTCTCTCTAGCACCCTGTGTC 304
|||||
QY 2660 TACGAGCAAAAGGGATAGGTTTGGAAAACCTATCATTTCTGCTCTTCTTAAGAGGAAA 2719
Db 305 TACGAGCAAAAGGGATAGGTTTGGAAAACCTATCATTTCTGCTCTTCTTAAGAGGAAA 364
|||||
QY 2720 CGAGCTCTGTAGAAAGCAAGACTTCTTTTCTGGCTCTTCTTCTCAACCTAA 2779
Db 365 CGAGCTCTGTAGAAAGCAAGACTTCTTTTCTGGCTCTTCTTCTCAACCTAA 434
|||||
QY 2780 AGCCAGAAAACCTTGCAGAGTATTGTTGGGGTGTGTTTATATTAGGCATTTGGGGGA 2839
Db 425 AGCCAGAAAACCTTGCAGAGTATTGTTGGGGTGTGTTTATATTAGGCATTTGGGGGA 484
|||||
QY 2840 TGGGGTGGAGGGGTATATAGTTTATAGGGTTTCTAAGAAATGCTAACAAAGCACTT 2899
Db 485 TGGGGTGGAGGGGTATATAGTTTATAGGGTTTCTAAGAAATGCTAACAAAGCACTT 544
|||||
QY 2900 TTGGACAATGCTATCCAGCAGGAAAAGGATAATACTGTTTAAACTCTTTC 2959
Db 545 TTGGACAATGCTATCCAGCAGGAAAAGGATAATACTGTTTAAACTCTTTC 604
|||||
QY 2960 TGGGGAATCCAATTATAGTTGCTTTGTTGTTTAAACCAAGAACAGCCAGGGTGTTCGC 3019
Db 605 TGGGGAATCCAATTATAGTTGCTTTGTTGTTTAAACCAAGAACAGCCAGGGTGTTCGC 664
|||||
QY 3020 CAGGGTAGAGTGTCTTAAAGATTGGTCCCTTGAATAATATGCTTCTGTATCAAAAGGTA 3079
Db 665 CAGGGTAGAGTGTCTTAAAGATTGGTCCCTTGAATAATATGCTTCTGTATCAAAAGGTA 724
|||||
QY 3080 CGTATGTGTGCACCAACAGGAGAACTTCTTTAAATTTCTTCTTCTTATTTTAAAC 3139
Db 725 GTATGTGGCGCAAAACAGCAGAACTTCTTTAAATTTCTTCTTCTTATTTTAAAC 784
|||||
QY 3140 AAATGGTGAAGATGGAGGATTACCTACAAATCAGACATGGCAAAACAATAATGGCTG-- 3197
Db 785 AATGGGTG-AAGATGGAGGATACCTACCAATCCAACTGGGCAACCAATATAGGGTGGT 843
|||||
QY 3198 TTTGCTTCCATAAACAAGTGCATTTTTTT 3226
Db 844 TTTGCTTCCATAAACAAGGCAATTTTTTT 872
|||||

Query Match	18.4%;	Score 696.4;	DB 7;	Length 852;
Best Local Similarity	92.6%;	Pred. No. 5.5e-158;		
Matches 787;	Conservative 0;	Mismatches 56;	Indels 7;	Gaps 5;
QY	2079	CTGGGAGAGAAAGATTCCGGGATTACTGATCTCCCAAGAGAGATCAGACATTACTTAT	2138	
Db	849	CGGGGAGAGAAAGATTCCGGGATTACTGATCTCCCAAGAGAGATCAGACATTACTTAT	790	
QY	2139	TGAATCAGCCTTTTGGAGCTGTTTGGCTCAGACATTTCCATCAGGTCAAACTCCTGA	2198	
Db	789	AGAATCAGCCTTTTGGAGCTGTTTGGCTCAGACATTTCCATCAGGTCAAACTCCTGA	730	
QY	2199	AGATAGTGTGCTCTGCAATCGACTGTCTCGATCGACTCAGTGCCTTCGTGGATT	2258	
Db	729	AGATAGTGTGCTCTGCAATCGACTGTCTCGATCGACTCAGTGCCTTCGTGGATT	670	
QY	2259	TGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAAATTTTCAGAGCTGAACCTTGA	2318	
Db	669	TGGGAGTGGCTCGACTCCATTAAAGACTTTTCTTAAATTTTCAGAGCTGAACCTTGA	610	
QY	2319	TATCCAGCCTTAGCCTGTGCTGTCAGACATGATGATCAGAGAGACATGGGTAAA	2378	
Db	609	TATCCAGCCTTAGCCTGTGCTGTCAGACATGATGATGATCAGAGAGACATGGGTAAA	550	
QY	2379	AGAACCAAGAGAGTCCGAGAGCTATGCAACAGATCAGAGAGTTTAAAGACCA	2438	
Db	549	AGAACCAAGAGAGTCCGAGAGCTATGCAACAGATCAGAGAGTTTAAAGACCA	490	
QY	2439	GAGTAAGGAGACAGCTCTGGAGCCACCGAGTCCAGGTCTCGGTGCCCTGTGAAC	2498	
Db	489	GAGTAAGGAGACAGCTCTGGAGCCACCGAGTCCAGGTCTCGGTGCCCTGTGAAC	430	
QY	2499	GAGGAAGATCTGCACCTGGGCTCCAGGCGATCTTCTACCTGAAGCTGGAAC	2558	
Db	429	GAGGAAGATCTGCACCTGGGCTCCAGGCGATCTTCTACCTGAAGCTGGAAC	370	
QY	2559	GTCTCCACCTCCATTTGACAGCTCTTCTGGACACCTTCTTATCAGGAGC	2618	
Db	369	GTCTCCACCTCCATTTGACAGCTCTTCTGGACACCTTCTTATCAGGAGC	310	
QY	2619	AG-TGGAGAGTGGAGTGGCTCTCTCTAGACCTGCTTGTCTACGAGAGGGA-TA	2676	
Db	309	AGCTGAGCAGGAGTGGCTCTCTCTAGACCTGCTTGTCTAGTACGAGGAGG	250	
QY	2677	GGTTTGGAACTTATCATTTCTGCTCTTCTTAAAGAG-AAAAGAGCTCTCTGTAGAAA	2735	
Db	249	GGTCTGGACACCTTATCTGCTCTTCTTAAAGAGAAAACAAGCTCTCTGTGAAA	190	
QY	2736	GCAAGACTTTCTTTT---TTTCTGCTCTTCTTCTTACACCTTAAAGCGAAGCTT	2792	
Db	189	TGAAGACTTTTCTTTTCTGCTCTTCTTCTTACATCTGAAAGCGAAGAACTT	130	
QY	2793	GCAGAGTATTTGTTGGGTTGTTTATTTATTTAGGCAATGGGGATGGGAGGG	2852	
Db	129	GCAAGTATTTGTTGGTGTGTTTATTTATTTAGG-TTTTGGGTTGGATGGGAGT	71	
QY	2853	GGTTATAGTCAAGGGTTTCTAAGAAATTTGCTAAAGAGCACTTTTGGCAATGCTA	2912	
Db	70	GGGTATAGTCAAGGGTTTCTAAGAAATTTGCTAAGAGCACTTTTGGCAATGCTA	11	
QY	2913	TCCAGCAGG 2922		
Db	10	TCCAGCAGG 1		
RESULT 6				
CB445292				
LOCUS	671 bp	mRNA	linear	EST 25-MAR-2003
DEFINITION	696544 MARC 6BOV Bos taurus	cdna 5',	mRNA sequence.	
ACCESSION	CB445292			
VERSION	CB445292.1	GI:29236177		
KEYWORDS	EST.			
SOURCE	Bos taurus (cow)			
ORGANISM	Bos taurus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.			
REFERENCE	1 (bases 1 to 671)			
AUTHORS	Smith, T.P.L., Roberts, A.J., Echternkamp, S.E., Chitko-McKown, C.G., Wray, J.E. and Keele, J.W.			
TITLE	A second set of bovine ESTs from pooled-tissue normalized libraries			
JOURNAL	Unpublished (2003)			
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross_match v0.990329. Plate: FQ8053 row: 1 column: 18 Seq primer: GTAATACGACTCACTATAGG.			
FEATURES	Location/Qualifiers			
source	1..671			
	/organism="Bos taurus"			
	/mol_type="mRNA"			
	/db_xref="taxon:9913"			
	/tissue_type="pooled"			
	/lab_host="DH10B"			
	/clone_lib="MARC 6BOV"			
	/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI; library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."			
ORIGIN				
Query Match	16.5%;	Score 627.8;	DB 6;	Length 671;
Best Local Similarity	96.0%;	Pred. No. 2.6e-141;		
Matches 644;	Conservative 0;	Mismatches 27;	Indels 0;	Gaps 0;
QY	1755	CCGATGTCTCAGTACTCTCGATTTTTCAGAGTGTCTCAGTCTCGAATGGTTTAAAGAGTTGT	1814	
Db	1	CCGATGTCTCAGTACTCTCGATTTTTCAGAGTGTCTCAGTCTCGAATGGTTTAAAGAGTTGT	60	
QY	1815	CCGTACAGATAGTCTGAAAGGAGAGAGGTCGTCTGCCCTTCCAAACCAAGAGCCCAT	1874	
Db	61	CCGTACAGATAGTCTGAAAGGAGAGAGGTCGTCTGCCCTTCCAAACCAAGAGCCCAT	120	
QY	1875	ACAAACAGGAACCTCTCAGCCCTCTCCACCTTCTCTCCAACTGCATCATGAATGCCCT	1934	
Db	121	ACAGCAGGAAGCTCTCAGCCCTCTCCACCTTCTCTCCAACTGCATCATGAATGCCCT	180	
QY	1935	TGTCGAGCTTTTAAACAGACTCAACACCCAGAGATCTTGATTTTCCAGATCTGTCTCCAC	1994	
Db	181	TGTCGAGCTTTTAAACAGACTCAACACCCAGAGATCTTGATTTTCCAGATCTGTCTCCAC	240	
QY	1995	TCACAGGCTGTCTGACGACAGATCTGAGCATGTGCAACAAATTTCTACAACTCTCTGAC	2054	
Db	241	TCACAGGCTGTCTGACGACAGATCTGAGCATGTGCAACAAATTTCTACAACTCTCTGAC	300	
QY	2055	AGCCTCCATTGATGATCCAGAGCTGGGAGAGAAAGATTCCGGGATTTTACTGATCTCCC	2114	
Db	301	AGCCTCCATTGATGATCCAGAGCTGGGAGAGAAAGATTCCGGGATTTTACTGATCTCCC	360	
QY	2115	CAAGAAGATCAGACATTACTTATTAAGATCAGCCCTTTTGGAGCTGTTTGTCTCAGACT	2174	
Db	361	CAAGAAGATCAGACATTACTTATTAAGATCAGCCCTTTTGGAGCTGTTTGTCTCAGACT	420	
QY	2175	TTCCATCAGGTCAAAACACTGCTGAAGATAAGTTTGTGTCTGCAATGGAATTGTCTCTGA	2234	
Db	421	TTCCATCAGGTCAAAACACTGCTGAAGATAAGTTTGTGTCTGCAATGGAATTGTCTCTGA	480	
QY	2235	TCGACTTTCAGTCCCTTCGTGGATTTCGGGAGTGGCTCCGACTCTATTAAAGACTTTTCC	2294	

481	TCGACTTCAGTGCCTTCGTGGATTTCGGGAGTGGCTCGACTCCATTTAAGACCTTTTCCTT	540
Db		
2295	AAATTTGCAGAGCGCTGAACCTTCATATCCAAGCCTTTAGCTGCTCTGTCAGCACTGAGCAT	2354
Qy		
541	AGTTTTCGAGCGCTGAACTTGTATTCNAGCCTTTAGCTGCTGTTCAGCACTGAGCAT	600
Db		
2355	GATCACAGAAGAATCATGGTTTAAAGAAACCAAAGAGAGTCGAAGAGCTATGCAACAGAT	2414
Qy		
601	GGTCA CAGAACGACATGGGTTAAAGAAACCAAAGAGAGTGGAGGAGCTATGCAACAGAT	660
Db		
2415	CACAAGCAGTT	2425
Qy		
661	CACAAGCAGCT	671
Db		

RESULT 7	CD521819	757 bp	mRNA	linear	EST 06-JUN-2003
LOCUS	CD521819				
DEFINITION	AGENCOURT 14360118 NIH MGC 191 Homo sapiens cDNA clone				
ACCESSION	IMAGE:30409856 5', mRNA sequence.				
VERSION	CD521819				
KEYWORDS	CD521819.1	GI:31453537			
	EST				

ORIGIN

Query Match	16.3%	Score	617;	DB	6;	Length	757;
Best Local Similarity	98.9%;	Pred. No.	1.2e-138;				
Marches	632;	Conservative	0;	Mismatches	5;	Indels	2;
				Gaps	1;		

2866	QY	GAGGGTTTTCTAAGAAATGCTTAA	CAAAAGCACATTTTGGACAAATGCTATCCACGACGAAA	2926
1	Db	GGGGTTTTCTAAGAAATGCTTAA	CABAGCACATTTTGGACAATGCTATCCACGACGAAA	60
2926	QY	AAAAAGGATAATATATACTGTGTTTT	AAAAAATCTTTTCTGGGGAATCCAAATATATAGTTCCTTTG	2985
61	Db	AAAAAGGATAATATATACTGTGTTTT	AAAAAATCTTTTCTGGGGAATCCAAATATATAGTTCCTTTG	120
2986	QY	TATTTAAAAACAAGAACAGCCAAAGGGTGTG	TCGCCAGGGTAGGATGTGTCTTTAAAGATTTG	3045
121	Db	TATTTAAAAACAAGAACAGCCAAAGGGTGTG	TCGCCAGGGTAGGATGTGTCTTTAAAGATTTG	180
3046	QY	GTCCCTTGAAAAATATGCTTCCGTATCA	AAAGTAGTATCGGTGCAACAAGGACAGAA	3105
181	Db	GTCCCTTGAAAAATATGCTTCCGTATCA	AAAGTAGTATCGGTGCAACAAGGACAGAA	240
3106	QY	CTTCCCTTTAAATTTCCCTCTCTCTCTTT	TATTTTAAACAAATGGTGAAATGAGGATTA	3165
241	Db	CTTCCCTTTAAATTTCCCTCTCTCTCTTT	TATTTTAAACAAATGGTGAAATGAGGATTA	300
3166	QY	ACAATTCAGACATGGCAAAACAAATAT	ATGCTGTGTTTCCATAAACAAGTGCATTTTT	3225
301	Db	ACAATTCAGACATGGCAAAACAAATAT	ATGCTGTGTTTCCATAAACAAGTGCATTTTT	360
3226	QY	TAAAGTGCTGTCTTACTAAGTCTGTGTT	TATTAATTAATCTCTCTTTATGTAATATA	3285
361	Db	TAAAGTGCTGTCTTACTAAGTCTGTGTT	TATTAATTAATCTCTCTTTATGTAATATA	420
3286	QY	AAGGAGCAGTCATGTTAGCAATATACG	TGTTAATATCCTAGCAGAGGCTGTGTTCA	3345
421	Db	AAGGAGCAGTCATGTTAGCAATATACG	TGTTAATATCCTAGCAGAGGCTGTGTTCA	480
3346	QY	TTCCCTGTGCGATCCCTTCTGAGTGATG	GGCCATCCCAAGACTTTTAGGCCATTTCTG	3405
481	Db	TTCCCTGTGCGATCCCTTCTGAGTGATG	GGCCATCCCAAGACTTTTAGGCCATTTCTG	540
3406	QY	AACCAGATCCCTGCCCTGACTGTCCAG	CTATCCTGAAAGTGGATCAGATTTATAA	3465
541	Db	AACCAGATCCCTGCCCTGACTGTCCAG	CTATCCTGAAAGTGGATCAGATTTATAA	600
3466	QY	TTACATGATAACTGTTTTT--GGTTGT	GTGTTCTATCAACCCC	3502
601	Db	TTACATGATAACTGTTTTTGTGTTG	GTGTTCTATCAAAACC	639

RESULT 8	CB465379	679 bp	mRNA	linear	EST-26-MAR-2003
LOCUS	726652	MARC 6BOV	Bos taurus	cdna 5', mRNA	sequence.
DEFINITION	CB465379				
ACCESSION	CB465379.1	GI:29271764			
VERSION	EST.				
KEYWORDS	Bos taurus	(cow)			
SOURCE	Bos taurus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos;				
REFERENCE	1 (bases 1 to 679)				
AUTHORS	Smith,T.P.L., Roberts,A.J., Ehternkamp,S.E., Chitko-McKown,C.G., Wray,J.E. and Keele,J.W.				
TITLE	A second set of bovine ESTs from pooled-tissue normalized libraries				
JOURNAL	Unpublished (2003)				
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross_match v0.990329. Plate: LAM8005 row: F. column: 20 Seq primer: GTAATACGACTCACTATAGG.				

FEATURES		Location/Qualifiers	
source	1..679		
	/organism="Bos taurus"		
	/mol_type="mRNA"		
	/db_xref="taxon:9913"		
	/tissue_type="pooled"		
	/lab_host="DH108"		
	/clone_lib="MARC 6BOV"		
	/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."		
ORIGIN			
Query Match			
Best Local Similarity 16.1%; Score 612.2; DB 6; Length 679;			
Matches 648; Conservative 0; Mismatches 23; Indels 2; Gaps 2;			
QY	1694	AAAATGCAAAATATGTTTGCCTGGCAATATAAACTGCCAGTAGACAAGACGTCGAA	1753
DB	7	AAAATGCAAAATATGTTTGCCTGGCAATATAAACTGCCCGTAGACAAGACGTCGAA	66
QY	1754	ACCGATGTCAGTACTGTCGATTTCAAGAGTGTCTCAGTGTTCGAATGCTAAAGAACTTG	1813
DB	67	ACCGATGTCAGTACTGTCGATTTCAAGAGTGTCTCAGTGTTCGAATGCTAAAGAGTTG	126
QY	1814	TCCGTACAGATAGTCTGAAAGGGAGGAGAGGTCTGCTGCTTCCAAACCAAGAGCCCAT	1873
DB	127	TCCGTACAGACAGTCTGAAAGGGAGGAGAGGTCTGCTGCTTCCAAACCAAGAGCCCAT	186
QY	1874	TACAACAGAACTTCTCAGCCCTCTCACCCTTCTCCTCCAACTGCATGATGAATGCC	1933
DB	187	TACAGCAGGAAGCCTCTCAGCCCTCTCACCCTTCTCCTCCGATCTGATGAATGCC	246
QY	1934	TTGTCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTTTCCAGATCTGTCCCA	1993
DB	247	TTGTCGAGCTTTAAACAGACTCAACGCCAGAGATCTTGATTTTCCAGATCTGTCCCA	306
QY	1994	CTGACGAGCTGCTCGAGGACACAGATGCTGAGCATGTGCAACAAATCTCAACCTCTCA	2053
DB	307	CTGACGAGCTGCTCGAGGACGAGTGTGAGCATGTACACAGTCTCAACCTTCTAA	366
QY	2054	CAGCTCCATGATGATTCAGAAAGCTGGGAGGAGAAAGATTCGGGATTTACTGATCTCC	2113
DB	367	CAGCTCCATGATGATTCAGAAAGCTGGGAGGAGAAAGATTCGGGATTTACTGATCTCC	426
QY	2114	CCAAAGAAGATCAGACATTAATTGATCAGCTTTTGGAGCTGTTGCTCAGAC	2173
DB	427	CCAAAGAAGATCAGACATTAATTGATCAGCTTTTGGAGCTGTTGTTCTCAGAC	486
QY	2174	TTTCCATCAGGTCAAAACACTGCTGAAGATAAG-TTTGTGTTCTGCAATGGAATTTGCTG	2232
DB	487	TTTCCATCAGGTCAAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGGAATTTGCTG	546
QY	2233	CATCGATTCAGTGCCTTCGTGGAATTT-GGGGAGTGGCTCGACTTAAAGACTTTTTC	2291
DB	547	CATCGATTCAGTGCCTTCGTGGAATTTGGGGAGTGGCTCGACTTAAAGACTTTTTC	606
QY	2292	CTTAAATTTGCAGAGCTGAACCTTGATATCCAAAGCTTAGCCTGCTGTCAGACTGAG	2351
DB	607	CTTAAATTTGCAGAGCTGAACCTTGATATCCAAAGCTTAGCCTGCTGTCAGACTGAG	666
QY	2352	CATGATCAGAGAA 2364	
DB	667	CATGGTCACAGAA 679	
RESULT 9			
CB447412			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
CB447412.1 GI:29253794			
CB447412 635 bp mRNA linear EST 26-MAR-2003			
701414 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.			
CB447412			
CB447412.1 GI:29253794			

KEYWORDS		EST.	
SOURCE		Bos taurus (cow)	
ORGANISM		Bos taurus	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.	
REFERENCE		1 (bases 1 to 635)	
AUTHORS		Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G., Wray,J.E. and Keele,J.W.	
TITLE		A second set of bovine ESTs from pooled-tissue normalized libraries	
JOURNAL		Unpublished (2003)	
COMMENT		Contact: Smith TPL	
		USDA, ARS, US Meat Animal Research Center	
		PO Box 166, Clay Center, NE 68933-0166, USA	
		Tel: 402 762 4366	
		Fax: 402 762 4390	
		Email: smith@email.marc.usda.gov	
		Single pass sequencing. Bases called with phred v0.020425.c and	
		trimmed with the aid of the trim_alt option. Vector identified with	
		cross_match v0.990329.	
		Plate: FQY8056 row: B column: 15	
		Seq primer: GTAATACGACTCACTATAGG.	
FEATURES		Location/Qualifiers	
source		1..635	
		/organism="Bos taurus"	
		/mol_type="mRNA"	
		/db_xref="taxon:9913"	
		/tissue_type="pooled"	
		/lab_host="DH108"	
		/clone_lib="MARC 6BOV"	
		/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."	
ORIGIN		Query Match	
		Best Local Similarity 15.7%; Score 597.2; DB 6; Length 635;	
		Matches 611; Conservative 0; Mismatches 24; Indels 0; Gaps 0;	
QY	1619	ACAACGCCCGCTGCCAGCACTACGGCGTCCGAAACCTTCGAGGGCTGCAAGGGCTTTTCA	1678
DB	1	ACAACGCCCGCTGCCAGCACTACGGGGTGGCGACCTCGAGGGCTGCAAGGGCTTTTCA	60
QY	1679	AGAGAACAGTGCAGAAAAATGCAAAATATGTTTGCCTGCGCAATAAAAACTGCCAGTAG	1738
DB	61	AGAGAACAGTGCAGAAAAATGCAAAATATGTTTGCCTGCGCAATAAAAACTGCCCGTAG	120
QY	1739	ACAAGAGAGCTGAAACCGATGTTCAGTACTGTCGATTTCCAGAGTGTCTCAGTGTGGA	1798
DB	121	ACAAGAGAGCTGAAACCGATGTTCAGTACTGTCGATTTCCAGAGTGTCTCAGTGTGGA	180
QY	1799	TGCTAAAGAGTGTTCGCTACAGATAGTCTGAAAGGAGGAGAGGTCTGCTGCTTCCA	1858
DB	181	TGCTTAAAGAGTGTTCGCTACAGAGTCTGAAAGGAGGAGAGGTCTGCTTCCA	240
QY	1859	AACCAAGAGCCCATTTACAAAGGAACTTCTTCAGCCCTCTCCACCTTCTCTCCCAATCT	1918
DB	241	AACCAAGAGCCCATTTACAGAGGAGCCCTTCAGCCCTCTTCACCTTCTCTCCGATCT	300
QY	1919	GCATGATGAATGCCCTTGTCCGAGCTTTAAACAGCTCAACACCCAGAGATCTTGATTT	1978
DB	301	GTATGATGAATGCCCTTGTCCGAGCTTTAAACAGACTCAACGCCCAGAGATCTTGATTT	360
QY	1979	CCAGATACCTGCTCCCACTGACAGGCTCTGCGAGGACAGATGCTGAGCATGTGCACAT	2038
DB	361	CCAGATACCTGCTCCCACTGACAGGCTCTGCGAGGACAGATGCTGAGCATGTGCACACAGT	420
QY	2039	TCTACAACTCTCTGACAGCTTCCATTGATGTATCCAGAAAGCTGGGAGAGAAAGATTC	2098
DB	421	TCTACAACTCTTACAGCTTCCATTGATGTATCCAGAAAGCTGGGAGAGAAAGATTC	480
QY	2099	GATTTACTGATCTCTCCCAAGAGAGATCAGACATTAATTGATCAGCCTTTTGGAGC	2158

Db 481 GATTACTGATCCCAAGAAGATCAGACATTACTTATAGATCAGCCTTTTGGAGC 540
 QY 2159 TGTGTGTCCTCAGATTTCCATCAGGTCAACACAGCTGCTGAAGATAAGTTTGTGTCTGCA 2218
 Db 541 TGTGTGTCCTCAGATTTCCATCAGGTCAACACAGCTGCTGAAGATAAGTTTGTGTCTGCA 600
 QY 2219 ATGGACTGCTCTGCATCGACTTCAGTGCCTTCGT 2253
 Db 601 ATGGACTGCTCTGCATCGACTTCAGTGCCTTCGT 635

RESULT 10
 BU436288 940 bp mRNA linear EST 29-NOV-2002
 LOCUS 603208116F1 CSEQRN11 Gallus gallus cDNA clone CHEST19303 5', mRNA
 DEFINITION
 ACCESSION BU436288
 VERSION BU436288.1 GI:25925599
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 940)
 AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 TITLE A comprehensive collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 22335534
 PUBMED 12445392

COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers

FEATURES

source
 1..940
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Layer and broiler"
 /db_xref="taxon:9031"
 /clone="CHEST18303"
 /sex="Male and female"
 /tissue_type="muscle"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector: pBluescript II KS(+); Site 1: EcoRI;
 Site 2: NotI; This normalized library was constructed from
 1 million independent clones. cDNA synthesis was initiated
 using an oligo(dT) primer, using methylated C in the first
 strand synthesis reaction. Following this first strand
 reaction, double-stranded cDNA was blunted, ligated to
 NotI adapters, digested with EcoRI, size-selected, and
 cloned into the NotI and EcoRI compatible sites of a
 custom modified MCS of the pBluescript (KS+) vector. The
 library was normalized in 2 rounds using conditions
 adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
 Ronaldo et al., Genome Research 6 (1996): 791, except that
 a significantly longer reannealing hybridization was
 used."

ORIGIN

Query Match 15.4%; Score 583; DB 5; Length 940;
 Best Local Similarity 81.8%; Pred. No. 2.3e-130;
 Matches 703; Conservative 0; Mismatches 140; Indels 16; Gaps 2;
 QY 1767 CTGTGATTTTCAAGAGTCTCTCAGTGTGGTAAGTGAAGAGTGTCCGTACAGATAG 1826

Db 2 CTGCGGTTTTTCAAGAGTGTCTCAGCGTGGCATGGTGAAGAAGTTGTCCGCAAGGACAG 61
 QY 1827 TCTGAAGAGGAGGAGAGGTGGTCTGCTTCCAAACCAAGAGAGCCATTACAAAGAGAAC 1886
 Db 62 CTTGAAGAGGAGAGAGGTGGTCTGCTTCCAAACCAAGAGAGCCCTCGCAGGAGAAC 121
 QY 1887 TTCTCAGCCCTCTCCACCTTCTCTCTCAATCTGATGATGAATGCCCTTGTCCGAGCTTT 1946
 Db 122 TTGCGAGCCCTCCCGGCTTCTCTCTCCCATCAGCATGATGAACGCCCTCTGTAGAGCTTT 181
 QY 1947 AACAGACTCAACACCCAGAGATCTTGATTATCCAGATATCTGCCCACTCACCAGGCTGC 2006
 Db 182 AACGACTCCAGCCCGAGGAGCTCGACTATTTCAGATATCTGTTCCACCGATCAGGCTGC 241
 QY 2007 TGCAGGCACAGATCTCTGAGCATGTGCAACAATTTCTACAACCTCTCTGACAGCCTCCATGA 2066
 Db 242 TGCAGGCACAGATGCAACAATTTCTACAACCTCTCTGACAGCCTCCATGA 301
 QY 2067 TGTATCCAGAGCTGGGCGAGAAAGATTCGGGATTTTACTGATCTCCCAAGAGATCA 2126
 Db 302 CATATCCAGAGCTGGGCGAGAAATTTCCAGGATTTTACTGACCTCCCGAAGAGATCA 361
 QY 2127 GACATTACTTATTAATCAGCCTTTTGGAGCTGTTTGTCTCTCAGATCTTCCATCAGGTC 2186
 Db 362 GACATTACTCATGAATCAGCTTTTGGAGCTGTTTGTACTAAGACTCTCCATCAGGTC 421
 QY 2187 AAACACTGCTGAAGATAAGTTTGTCTGCAATGGAATCTGCTCTGATGATGATCAAGAG 2246
 Db 422 TGATACTGCTCAGGATAAGTTTGTATTTCTCAATGGACTTGTGTCTCATAGACTTCAGTG 481
 QY 2247 CTTCTGGGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAAATTTGCAGAG 2306
 Db 482 CTTCTGGGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAAACTTTAAAGAG 541
 QY 2307 CTTGAACCTTGATATCAAGCCTTAGCCTGCTCTCAGCACTGAGCATGATCAAGAGAG 2366
 Db 542 CTTGAACCTTGATATCAAGCCTTAGCCTGCTCTCAGCACTGAGCATGATCAAGAGAG 601
 QY 2367 ACATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAAGAGTGAAGAGTGAAGAGCTTT 2426
 Db 602 ACATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAAGAGTGAAGAGTGAAGAGCTTT 661
 QY 2427 AAAGAGCA-----CCAGAGTAAGGAGCAGAGGCTCTGGAGCCCAAGCGAGTC 2471
 Db 662 GAAGATCATTAACTTTAGTTGCCAACAAGAGCAGCGCTCGAGTCGCGAGAGGCC 721
 QY 2472 CAAGGTCTCTGGTGCCTCTGGTAGAACTGAGAAAGATCTGCACCTTGGGCTCCAGCGCAT 2531
 Db 722 TAAGTACTGGGTGTTCTTGTCTG-CTTCTGCACTCTGGGACTGCGAGCGCAT 780
 QY 2532 CTTTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCTTCAATATTCGACAGCTCTTCT 2591
 Db 781 CTTTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCTTCAATATTCGACAGCTCTTCT 840
 QY 2592 GCACACCTTACCTTTCTTAA 2610
 Db 841 GGACAACTTTGCTTCTGA 859

RESULT 11

CD638747 635 bp mRNA linear EST 17-JUN-2003
 LOCUS AGENCOURT_14534944 NIH_MGC_191 Homo sapiens cDNA clone
 DEFINITION IMAGE:30417493 5', mRNA sequence.
 ACCESSION CD638747
 VERSION CD638747.1 GI:31804872
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 635)

SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 802)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
The following repetitive elements were found in this cDNA
sequence: 86-116, >GC-rich#low_complexity
Seq primer: pyx-5.

FEATURES
source
1..802
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30643330"
/tissue_type="Upper Head"
/dev_stage="9.5-10.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_RN0"
/note="Organ: Head; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAAGTGAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN
Query Match 14.6%; Score 555.4; DB 7; Length 802;
Best Local Similarity 85.6%; Pred. No. 1.2e-123;
Matches 688; Conservative 0; Mismatches 104; Indels 12; Gaps 6;
QY 148 ACACAGCGCGCACAGAGCTCCGCGACACACTCTCCCGCGCGCTACACCCCT 207
DB 2 ACAGCGCGCTACAGCTCCGCGACACACTCTCTCCCGCGCGCTACACCCCT 61
QY 208 CTTCGCCCTGAGCCCTTCCCGGT-----GACGCGCGCGCGCGAGCTGACGCCCTCCCG 262
DB 62 CTCACCGCGCTCCTCGCAGGTGTCGCGCGCGCGCGCGAGCGCGCCCTCCAG 121
QY 263 GGTCTACTTTGCAAGCTGACGCTGCGCGAGTGGCGGTGGAGTGGAGACGCGCGCG 322
DB 122 GGCTCACTTTGCAACGCTGACAGAGCGCGAGTGTCCGTGGAGTGGGAACGTGGCGAC 181
QY 323 ATCTCTCCCTCGTTCACAGCCCAAGCCAGGACGCGCGGAACTCTCGGCTGTCTCT 382
DB 182 ATCTTACCCCTCGTTCGAGCCGAGACTGGATGCTTGGAACTCTCGGCGCGCTCT 241
QY 383 CCCATGAGTGGGATCGCAGCATCCCCCAACAGCGCTCTACCGCTCTCGGAGCGCGCTG 442
DB 242 CCCATGAGTGGGATCGCAGCATCCCCCGCGCGCGCTCTACCGCTCTGGAGCGCGCTG 301
QY 443 GCTTGTACACCGAGCCCTTCGGGGACAGCAGCTGTGACTCCCCCGCAGTGCAGATTTCG 502

DB 302 GTTTGTGACCGAGCCCTTCGGGACACAGCTGTGACTCTCCCCCAATTCAGATTTCG 361
QY 503 GGACAGCTCTCTAGAAACTCGCTCTAAAGACGGAACCGGCACAGCACTCAAGGCCCACTG 562
DB 362 GGGTCGCTCTCTAGAAACTCGCTCTAAAGACGGAACCGGCACAGCACTCAAGGCCCACTG 421
QY 563 CGAGAGAGGAGCGCGCGGCAAGCGCGGCGCTGAGCTGGAGCCCTTAGGGGTGCCGGGCA 622
DB 422 CGGAGAGGCGGCGCGGCAAGCGCGGCGCTGAGCTGGAGCCCTTAGGGGTGCCGGGCA 481
QY 623 GCA---CTGCGCGCGCTTCGCTCGCGGAGCTCGCTCTCTACACTCTACGCTCCG 679
DB 482 GGACAGCGGCTGCTGCTTCCCTATCCGAGCTTCCCGCTCTCTACACTCTACGCTCCG 541
QY 680 CTGGAGAGACCCCGAGCCCACTTACGCGGCGAAGATACCTCCAGATATGCCCTCGG 739
DB 542 CTGGAGAGA-CCCGAGCCCACTTACGCGGCGAAGATACCTCCAGATATGCCCTCGG 600
QY 740 TCAAGCCCAATATAGCCTTCCCTCCAGGTTCCGTTATGGGGGCGAGACATACAGCT 799
DB 601 TGCNAAGCCCAATATAGCCTTCCCTCCGCGGTCCACTTACGCCACGCACTTATGGCT 660
QY 800 CGGATATACACCGGAGATCATGAACCGGAGCTTACACCAAGCTGACCATGGACCTTGGCA 859
DB 661 CGGAATA-ACCACAGAAATCATGAA-CCGACTTACACCAAGCTGACCATGGACCTCGGTA 718
QY 860 GCACCTGAGATACGGCTACAGCCACCACTGCTCCCTGCCAGCATCAGTACCTTCGTGGAGG 919
DB 719 GCACGGGATCATGNNACCGCCACTACATCCCTGCCAGCTTCAGTACCTTCATGGAGG 778
QY 920 GCTACTCGAGCAACTACGAATCA 943
DB 779 GCTAC-CCAGCAGCTCGCAACTCA 801
CB445564 644 bp mRNA linear EST 26-MAR-2003
696928 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
CB445564
CB445564.1 GI:29251946
EST.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
1 (bases 1 to 644)
Smith, T.P.L., Roberts, A.J., Ehternkamp, S.E., Chitko-McKown, C.G.,
Wray, J.E. and Keele, J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries
Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: PQY8053 row: I column: 18
Seq primer: TAGAAGGCACAGTCGAGG.
FEATURES
Location/Qualifiers
1..644
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pCDNA3.1; Site_1: EcoRI; Site_2: NotI;

Library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."

ORIGIN		Query Match	14.6%; Score 552.6; DB 6; Length 644;
		Best Local Similarity	92.2%; Pred. No. 5.3e-123;
		Mismatches	593; Conservative 0; Mismatches 49; Indels 1; Gaps 1;
Qy	2070	ATCCAGAGCTGGCGAGAAAGATTCCGGGATTTACTGATCTCCCAAGAGATCGAC	2129
Db	644	ATCCAGAGCTGGCGAGAAAGATTCCGGGATTTACTGATCTCCCAAGAGATCGAC	585
Qy	2130	ATTACTTATTGAATCAGCCCTTTTGGAGCTGTTTGTCTCAGACTTTCATCAGGTCAA	2189
Db	584	ATTACTTATTGAATCAGCCCTTTTGGAGCTGTTTGTCTCAGACTTTCATCAGGTCAA	525
Qy	2190	CACCTGCTGAAGATAAGTTTGTGTTTCAATGGACTTGTCTGCAATCGACTTCAGTGCCT	2249
Db	524	CACCTGCTGAAGATAAGTTTGTGTTTCAATGGACTTGTCTGCAATCGACTTCAGTGCCT	465
Qy	2250	TCGTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCCTTAAATTTTCAGAGCCT	2309
Db	464	TCGTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCCTTAAATTTTCAGAGCCT	405
Qy	2310	GAACCTTGATATCCAAAGCCTTAGCCTGCTGTCAGCACTGAGCATGATCACAGAAAGACA	2369
Db	404	GAACCTTGATATCCAAAGCCTTAGCCTGCTGTCAGCACTGAGCATGATCACAGAAAGACA	345
Qy	2370	TGGGTTAAAGAACCAAGAGAGTCGAAGAGCTATGCAACAAGATCAAGCAGTTTAAA	2429
Db	344	TGGGTTAAAGAACCAAGAGAGTCGAAGAGCTATGCAACAAGATCAAGCAGTTTAAA	285
Qy	2430	AGACCCAGAGTAAGGACAGCTCTGGAGCCCAAGGCTCCTGAGGTCCTGGTGCCCT	2489
Db	284	AGACCCAGAGTAAGGACAGCTCTGGAGCCCAAGGCTCCTGAGGTCCTGGTGCCCT	225
Qy	2490	GGTAGAAGTGAAGAGATCTGCACCTTGGGCTTCCAGCGCATCTTCTACCTGAAGCTGA	2549
Db	224	GGTAGAAGTGAAGAGATCTGCACCTTGGGCTTCCAGCGCATCTTCTACCTGAAGCTGA	165
Qy	2550	AGACTTGGTCTCCACCTTCCATCATTTGACAGCTCTTCCCTGGACACCTTACCTTTCTA	2609
Db	164	AGACTTGGTCTCCACCTTCCATCATTTGACAGCTCTTCCCTGGACACCTTACCTTTCTG	105
Qy	2610	ATCAGGAGCAG-TGGAGCAGTGAAGCTGCCTCTCTCTAGCACCTGCTTGTACGAGCA	2668
Db	104	ATCAGGAGCAG-TGGAGCAGTGAAGCTGCCTCTCTCTAGCACCTGCTTGTACGAGCA	45
Qy	2669	AAGGATAGGTTTGGAAACCTATCATTTCTGCTTCTTAA	2711
Db	44	GAGGGGGGGTCTGGACACCTACCTTTTCTGCTTCTTCAAAA	2

Search completed: March 21, 2005, 20:40:37
Job time : 11552 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2005, 17:36:21 ; Search time 170 Seconds
(without alignments)
1424.188 Million cell updates/sec

Title: US-10-608-863-2
Perfect score: 3337
Sequence: 1 MFCVQAQVSPSPGSSYAAQ.....EDLVSPPSIIDKFLDTLPF 626

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A: Geneseq_16Dec04:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3337	100.0	626	AD113008	AD113008 Human NOR
2	3333	99.9	626	ADP45995	ADP45995 Human Pro
3	3333	99.9	626	ADP28879	ADP28879 Human nor
4	3333	99.9	626	ADP05709	ADP05709 Human nuc
5	3316	99.4	626	ADL57171	ADL57171 Human NOV
6	3301	98.9	626	AAW16398	AAW16398 Human neu
7	3299.5	98.9	625	AAU96995	AAU96995 Human nuc
8	3299.5	98.9	625	ADF28878	ADF28878 Human nor
9	3133.5	93.9	643	ADF28862	ADF28862 Pig nor-1
10	3123	93.6	587	ADL83262	ADL83262 Human PRO
11	3123	93.6	587	ADR14063	ADR14063 Human NF-
12	3074	92.1	628	AA92057	AA92057 Apoptotic
13	3074	92.1	628	AB998438	AB998438 Murine Ne
14	3074	92.1	628	ADP45993	ADP45993 Rat Prote
15	3074	92.1	628	ADF28880	ADF28880 Rat nor-1
16	3054.5	91.5	627	AB998437	AB998437 Murine Ne
17	2099.5	62.9	446	ADF28881	ADF28881 Pig nor-1
18	2036.5	61.0	429	ADP05711	ADP05711 Mouse nuc
19	1751	52.5	598	AA48631	AA48631 Sequence
20	1751	52.5	598	AAU09156	AAU09156 Human NOT
21	1751	52.5	598	AA835156	AA835156 Human nuc
22	1751	52.5	598	ABR42448	ABR42448 Human Nur
23	1751	52.5	598	ADP84066	ADP84066 Human NUR
24	1751	52.5	598	ADP84062	ADP84062 Human NUR
25	1751	52.5	598	ADP84067	ADP84067 Human NUR

26	1751	52.5	598	7	ADB84068	ADB84068 Human NUR
27	1751	52.5	598	7	ADB83989	ADB83989 Human NUR
28	1751	52.5	598	7	ADB84070	ADB84070 Human Pro
29	1751	52.5	598	7	ADE56986	ADE56986 Human Pro
30	1751	52.5	598	7	ADP25771	ADP25771 Human tra
31	1751	52.5	598	8	ADH75170	ADH75170 Human tra
32	1751	52.5	598	8	ADP05705	ADP05705 Human nuc
33	1751	52.5	598	8	ADP23740	ADP23740 PRO polyp
34	1746	52.3	598	5	AAU96994	AAU96994 Mouse nuc
35	1746	52.3	598	7	ADP84074	ADP84074 Murine NU
36	1746	52.3	598	7	ADP84065	ADP84065 Murine NU
37	1746	52.3	598	7	ADP84064	ADP84064 Murine NU
38	1746	52.3	598	7	ADP84061	ADP84061 Murine NU
39	1746	52.3	598	8	ADL70235	ADL70235 Mouse Nur
40	1746	52.3	598	8	ADP05707	ADP05707 Mouse nuc
41	1736	52.0	598	7	ADP84069	ADP84069 Rat NURR1
42	1736	52.0	598	7	ADE56984	ADE56984 Rat Prote
43	1732	51.9	607	4	AAU09018	AAU09018 Mouse Flg
44	1684.5	50.5	597	6	ABR42447	ABR42447 Rat Nurrl
45	1684.5	50.5	597	7	ADP28899	ADP28899 Rat nur77

ALIGNMENTS

RESULT 1
AD113008
ID AD113008 standard; protein; 626 AA.

XX AC AD113008;

XX DT 22-APR-2004 (first entry)

XX DE Human NOR-1 (MINOR) protein sequence SeqID 2.

XX KW human; receptor; allergic disease; NOR-1; MINOR; eosinophil;

XX KW atopic dermatitis; antiallergic; antiinflammatory; dermatological.

XX OS Homo sapiens.

XX PN WO2004003198-A1.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-JP008199.

XX PR 27-JUN-2002; 2002JP-00188490.

XX PA (GENO-) GENOX RES INC.

XX PI (NIGE-) JAPAN GEN AGENCY NATION.

XX PI Hashida R, Kagaya S, Yayoi Y, Sugita Y, Saito H;

XX DR WPI; 2004-083057/08.

XX DR N-PSDB; AD113007.

XX PT Examining allergic diseases e.g. atopic dermatitis by differential display based on gene expression of NOR-1 receptor protein, also applicable in screening compounds for treatment of allergic diseases.

XX PS Example 1; SEQ ID NO 2; 155pp; Japanese.

XX CC This invention relates to a novel method for examining allergic diseases that comprises comparing the expression levels of a gene encoding the NOR-1 receptor protein between patients and healthy individuals.

XX CC Specifically, the NOR-1 gene, also referred to as MINOR, is expressed in the specialist white blood cells known as eosinophils and is involved in mediating an allergic reaction. The present invention describes a differential display method that can identify the expression level of this gene in order to identify its usefulness in diagnosing allergic diseases such as atopic dermatitis. Furthermore, compositions can also be used to screen compounds for the treatment of allergic diseases.

XX CC Accordingly, they exhibit various activities including antiallergic,

QY 301 COHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNRQYCRFQKCLSVGMVKE 360
DB 301 COHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNRQYCRFQKCLSVGMVKE 360
QY 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPPICMNVALVRALTDSTPRDLDSRYC 420
DB 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPPICMNVALVRALTDSTPRDLDSRYC 420
QY 421 PTDQAAAGTDAEHVQOYFNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 480
DB 421 PTDQAAAGTDAEHVQOYFNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 480
QY 481 RLSIRSNTEADKVFVFCNGVLVRLQCLRGFGWLSIKDFSLNLSLNLDIOALACLSAL 540
DB 481 RLSIRSNTEADKVFVFCNGVLVRLQCLRGFGWLSIKDFSLNLSLNLDIOALACLSAL 540
QY 541 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKTCTIGLOR 600
DB 541 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKTCTIGLOR 600
QY 601 IFYKLEDLVSPPSIIDKFLDLPF 626
DB 601 IFYKLEDLVSPPSIIDKFLDLPF 626

RESULT 3

ID ADF28879 standard; protein; 626 AA.

AC ADF28879;

DT 12-FEB-2004 (first entry)

DE Human nor-1 nuclear receptor polypeptide Q92570.

KW Human; nor-1; nuclear receptor; receptor; leukaemia; gene therapy;
KW cytostatic; haematopoietic cell.

OS Homo sapiens.

PN WO2003088812-A2.

PD 30-OCT-2003.

PF 15-APR-2003; 2003WO-US011804.

PR 17-APR-2002; 2002US-0373238P.

PA (BAYU) BAYLOR COLLEGE MEDICINE.

PA (UNIW) UNIV WASHINGTON.

PI Mulligan SE, Conneely OM, Milbrandt J;

PI WPI: 2003-854017/79.

DR GENBANK; Q92570.

PT Inhibiting proliferation of a hematopoietic cell, useful for treating or
PT preventing leukemia, comprises modulating the level or activity of nor-1
PT and/or nur77 nuclear receptor.

PS Disclosure; SEQ ID NO 14; 101pp; English.

XX The present sequence is that of human nor-1 nuclear receptor polypeptide
CC Q92570. Nor-1 has been identified as a molecular target for therapeutic
CC intervention in the treatment of myeloid leukaemia. The invention
CC provides methods of inhibiting the proliferation of a haematopoietic stem
CC cell or a haematopoietic myeloid cell by modulating the level of a nor-1
CC and/or nur77 nuclear receptor. This involves increasing the level of the
CC receptor polypeptide or polynucleotide, e.g. by administration of a
CC vector comprising the polynucleotide. A claimed method of treating
CC leukaemia comprises modulating a nor-1 and/or nur77 nuclear receptor in a
CC haematopoietic stem cell or myeloid cell. Also claimed are methods of

CC identifying an upregulator of expression of nor-1 and/or nur77, of
CC identifying a compound for the treatment of leukaemia, and of screening
CC for a compound for treatment of leukaemia, and a mouse model for
CC leukaemia comprising a mouse having defective nor-1 and nur77 nucleic
CC acid sequences.

XX SQ Sequence 626 AA;

Query Match 99.9%; Score 3333; DB 7; Length 626;

Best Local Similarity 99.8%; Pred. No. 3.1e-222;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPCVQAQYSPSPGSSYAAQTYSSVYTEIMNPDTYKLTMDLGSITEITATATSLPSIST 60

DB 1 MPCVQAQYSPSPGSSYAAQTYSSVYTEIMNPDTYKLTMDLGSITEITATATSLPSIST 60

QY 61 FVEGYSSNYELKPSVYQMRPLIKVEEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHH 120

DB 61 FVEGYSSNYELKPSVYQMRPLIKVEEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHH 120

QY 121 SSPDEVLPSTSMYFKQSPPTPTTAPPPQAGALWDEALPSAPGCIAPGILLDPPMKAV 180

DB 121 SSPDEVLPSTSMYFKQSPPTPTTAPPPQAGALWDEALPSAPGCIAPGILLDPPMKAV 180

QY 181 PTVAGARFPLFHFKPSPPHPPAPSPAGGHHLGYDPTAAALSLPLGAAAAAGSQAALBS 240

DB 181 PTVAGARFPLFHFKPSPPHPPAPSPAGGHHLGYDPTAAALSLPLGAAAAAGSQAALBS 240

QY 241 HPYGLPLAKRAAPLAPPLGLTPTSTASSLLGESPSLPSPPSRSSSSSGTGCAVCGDNAA 300

DB 241 HPYGLPLAKRAAPLAPPLGLTPTSTASSLLGESPSLPSPPSRSSSSSGTGCAVCGDNAA 300

QY 301 COHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNRQYCRFQKCLSVGMVKE 360

DB 301 COHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNRQYCRFQKCLSVGMVKE 360

QY 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPPICMNVALVRALTDSTPRDLDSRYC 420

DB 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPPICMNVALVRALTDSTPRDLDSRYC 420

QY 421 PTDQAAAGTDAEHVQOYFNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 480

DB 421 PTDQAAAGTDAEHVQOYFNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 480

QY 481 RLSIRSNTEADKVFVFCNGVLVRLQCLRGFGWLSIKDFSLNLSLNLDIOALACLSAL 540

DB 481 RLSIRSNTEADKVFVFCNGVLVRLQCLRGFGWLSIKDFSLNLSLNLDIOALACLSAL 540

QY 541 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKTCTIGLOR 600

DB 541 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKTCTIGLOR 600

QY 601 IFYKLEDLVSPPSIIDKFLDLPF 626

DB 601 IFYKLEDLVSPPSIIDKFLDLPF 626

RESULT 4

ADP05709

ID ADP05709 standard; protein; 626 AA.

XX AC ADP05709;

XX DT 26-AUG-2004 (first entry)

XX DE Human nuclear receptor protein SeqID83.

XX disease risk; disorder risk; mutation; polymorphism;
XX nuclear receptor protein; antibacterial; antithyroid; cardiovascular-Gen;
XX cytostatic; dermatological; eating-Disorders-Gen; gastrointestinal-Gen;
XX gynaecological; hepatotropic; immunosuppressive; muscular-Gen;
XX nephrotropic; osteopathic; virucide; adrenal gland; colon;
XX cardiovascular; intestine; kidney; liver; lung; muscular; ovary; blood;

Zhong M, Guo X, Anderson DW, Ort T, Padigaru M, Rieger DK;

WPI; 2004-315567/29.

N-PSDB; ADL57170;

New isolated NOVX polypeptides and polynucleotides, useful for preventing, diagnosing or treating NOVX-associated disorders, e.g. osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease, asthma or infections.

Claim 1; SEQ ID NO 116; 214pp; English.

The invention relates to a novel isolated polypeptide (NOVX) comprising a mature form of any of the 37 amino acid sequences fully defined in the specification. A polypeptide of the invention has antidiabetic, antineurotic, cardiac, hypotensive, antiarteriosclerotic, anorectic, virucide, antibacterial, fungicide, protozoacide, neurotropic, neuroprotective, antiparkinsonian, anticonvulsant, osteopathic, antiarthritic, antiinflammatory, dermatological, antiasthmatic, and antilipaeamic activity. A polynucleotide of the invention may have a use in gene therapy. The polypeptides, nucleic acid molecules and antibodies are useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOVX-associated disorder. The nucleic acid molecules, polypeptides and antibodies are useful for treating, preventing or diagnosing diseases such as metabolic disorders, diabetes, obesity, infectious diseases (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular diseases (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune disorders (osteoarthritis), haematopoietic disorders, inflammatory skin disorders, asthma, and various dyslipidaemias. The nucleic acids and polypeptides may also be used as targets for the identification of small molecules that modulate or inhibit e.g. neurogenesis, cell differentiation, cell proliferation, haematopoiesis, wound healing and angiogenesis, in gene therapy, in generation of antibodies that bind immunospecifically to NOVX substances for use in therapeutic or diagnostic methods. The nucleic acids are further used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. The NOVX polypeptides of the invention show homology to certain known human proteins: NOV1a-1t show homology to fibroblast growth factor receptor 4 (FGFR4); NOV2a shows homology to complement factor 1 precursor; NOV3a shows homology to matrix metalloproteinase-15 precursor; NOV4a shows homology to MDG3; NOV5a-3c show homology to T-lymphocyte surface antigen Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21 (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like polypeptide variant; NOV8a-8g show homology to antileukoprotease 1 precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology to nuclear hormone receptor NHR-1; NOV11a-11j show homology to transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin-dynorphin precursor. The present sequence represents a NOVX polypeptide of the invention.

Sequence 626 AA;

Every Match 99.4%; Score 3316; DB 8; Length 626;
 1st Local Similarity 99.4%; Pred. No. 4.6e-221;
 Matches 622; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 MPCVQAQYSPSPPGSSYAAQYSSSEYTTTEIMNPDYTKLTMDLGSTEITATATSLPSIST 600

1 MPCVQAOYSPSPGSSVAAOTYSSEYTTETMNPDYTKI.TMDI.GSSTEITATATTSI.PSIST 60

61 FVEGYSSNVEI.KPSCVVOMOBPI.IKVFEGRAPSVHHHHHHHHHHH000H00RSTPBA 120

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

.....

I 81 FIVAGARFFLFHFAPSPFAGGHHLDGIDFIAAAALSLPLGAAAAAGSQAAALLES 240

Qy	241	HPYGLPLAKRAAPLAFPLGLTPSPASSLLGESPSLPPSRSSSSGEGTCAVCGDNAA	300
Db	241	HPYGLPLAKRAAPLAFPLGLTPSPASSLLGESPSLPPSRSSSSGEGTCAVCGDNAA	300
Qy	301	COHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVPDKRRNRNCQYCRFKQKCLSVGMVKE	360
Db	301	COHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVPDKRRNRNCQYCRFKQKCLSVGMVKE	360
Qy	361	VVRTDSLKRRGRPLSPKPKSLQOEPSOPSPSPPICMNNALVRALTDTSTPRDLDSRYC	420
Db	361	VVRTDSLKRRGRPLSPKPKSLQOEPSOPSPSPPICMNNALVRALTDTSTPRDLDSRYC	420
Qy	421	PTDQAAAGTDAEHVQQVFYNLILTASIDVSRSWAEKIPGFTDLPKEQDTLLIESAFLEFLVL	480
Db	421	PTDQAAAGTDAEHVQQVFYNLILTASIDVSRSWAERIPGFTDLPKEQDTLLIESAFLEFLVL	480
Qy	481	RLSIRSNTAEDKFVFCNGVLVLRQLCGFGEWLDSIKDFSLNQSLLNDIQALACLAL	540
Db	481	RLSIRSNTAEDKFVFCNGVLVLRQLCGFGEWLDSIKDFSLNQSLLNDIQALACLAL	540
Qy	541	SMITERHGLKEPKRVEELCNKITSCLKDHQSKGQALEPTESKVLGALVELRKICTLGLQR	600
Db	541	SMITERHGLKEPKRVEELCNKITSCLKDHQSKGQALEPTESKVLGALVELRKICTLGLQR	600
Qy	601	IFYUKLEDLVSPPSIIDKFLDITLPPF	626
Db	601	IFYUKLEDLVSPPSIIDKFLDITLPPF	626

RESULT 6

AAW16398

AAW16398
ID AAW16398 standard; protein; 626 AA.

XX
XX

AC AAW16398;

XX

DT 08-SEP-1997 (first entry)

XX

DE Human neuron-derived

XX

KW Human; neuron-derived orphan receptor 1

KW NOR-1; DNA binding domain
 KW Parkinson's disease

KW parkinson's disease.
v v

XX Homo sapiens

OS Homo sapiens
yy

XX	Key	Location/Qual
FH		

FH	Key	Location/Qual
FT	Domain	291 377

FT	Domain	291:
FT		/not

XX
FI
2017/

JP09084585-A.

FN		UFG08043859-A.	
XX	-		
PD	31-MAR-1997.		
XX			
XX		95JP-00242909.	
Pf	21-SEP-1995;		
XX			
PR	21-SEP-1995;	95JP-00242909.	
XX			
PA	(TERU) TERUMO CORP.		
XX			
XX			
DR	WPI; 1997-253003/23.		
N-	PSDB; AAT73334.		
XX			
PT	Human neuron-derived orphan nuclear receptor gene - useful for diagnosing		
PT	brain diseases such as Alzheimer's or Parkinson's disease or		
PT	neuroblastoma.		

PS Claim 1: Page 5-8: 8pp: Japanese.

XX
 Page 2-0, opp, vapores: 22

THIS IS THE AMINO ACID SEQUENCE OF
nuclear receptor. The protein has

CC nuclear receptor: the protein has
CC receptor. The protein contains se

CC domains. The nucleic acid and protein can be used to diagnose brain
CC neuronal diseases, e.g. Alzheimer's disease, Parkinson's disease.
CC Further, the genetic diseases can be treated using the receptor as it
CC inhibits expression of a gene by binding to it
XX
SQ Sequence 626 AA;

Query Match 98.9%; Score 3301; DB 2; Length 626;
Best Local Similarity 98.9%; Pred. No. 5e-220;
Matches 619; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MPCVQAQYSPSPGSSYAAQTSYSSYTTIIMNPDYTKLTMDLGSSTIATATTSLSIST 60
Db 1 MPCVQAQYSPSPGSSYAAQTSYSSYTTIIMNPDYTKLTMDLGSSTIATATTSLSIST 60

QY 61 FVEGYSSNYELKPCVYQWQRLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120
Db 61 FVEGYSSNYELKPCVYQWQRLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120

QY 121 SPDEVLSTSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180
Db 121 SPDEVLSTSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180

QY 181 PTVAGARFPLFHFKSPSPHPAPSPAGGHHGLGYDPTAAAGLSLAVGAAAAGSQAAL 240
Db 181 PTVAGARFPLFHFKSPSPHPAPSPAGGHHGLGYDPTAAAGLSLAVGAAAAGSQAAL 240

QY 241 HPYGLPLAKRAAPLAPFPPLGLTPSTASSLLGESPLSPSPSSSGEGTCAVGDNA 300
Db 241 HPYGLPLAKRAAPLAPFPPLGLTPSTASSLLGESPLSPSPSSSGEGTCAVGDNA 300

QY 301 COHYGVRTCEGCKGFFKRTVQKNKVCVLANKNCPVDKRRNRQYCRFQKCLSVGMWKE 360
Db 301 COHYGVRTCEGCKGFFKRTVQKNKVCVLANKNCPVDKRRNRQYCRFQKCLSVGMWKE 360

QY 361 VRTDSLKGRRLPSPKSPSQEPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420
Db 361 VRTDSLKGRRLPSPKSPSQEPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420

QY 421 PTQAAAGTDAEHVQFNYLLTASIDVSRSWAEKIRGFTDLKEDOTLLIESAFLEFVL 480
Db 421 PTQAAAGTDAEHVQFNYLLTASIDVSRSWAEKIRGFTDLKEDOTLLIESAFLEFVL 480

QY 481 RLSIRNTAEDKVFVNCGLVHLRLQCLRGFGEWLDISKDFSLNQLSINLDIQALACL 540
Db 481 RLSIRNTAEDKVFVNCGLVHLRLQCLRGFGEWLDISKDFSLNQLSINLDIQALACL 540

QY 541 SMITERHGLKEPKRVEELCNKITSLSKHQSKQALPEPTESKVLGALVELRKICTLGL 600
Db 541 SMITERHGLKEPKRVEELCNKITSLSKHQSKQALPEPTESKVLGALVELRKICTLGL 600

QY 601 IFYKLEDLVSPSPSIIDKFLDLPF 626
Db 601 IFYKLEDLVSPSPSIIDKFLDLPF 626

RESULT 7
AAU96995
ID AAU96995 standard; protein; 625 AA.
AC AAU96995;
XX
XX 30-JUL-2002 (first entry)
DT Human nuclear receptor NOR1 protein sequence.
DE
XX Human; nuclear receptor; NURR; inflammatory immune disease; arthritis;
KW corticotropin releasing hormone; receptor; CRH; rheumatoid arthritis;
KW chronic inflammatory joint disease; psoriatic arthritis; thyroiditis;
KW sarcoid arthritis; ulcerative colitis; NOR1.
XX Homo sapiens.
OS
XX

PN WO200187923-A1.
XX
PD 22-NOV-2001.
XX
PF 11-MAY-2001; 2001WO-US015311.
XX
PR 12-MAY-2000; 2000US-0203645P.
XX
XX (BAYU) BAYLOR COLLEGE MEDICINE.
PA Murphy E, Conneely OM, Fitzgerald O, Bresnahan B;
PI WPI; 2002-075311/10.
XX
DR
XX
XX
PT Treating inflammatory immune disease such as arthritis, comprises
PT suppressing expression level of NURR subfamily of nuclear transcription
PT factors, or corticotropin releasing hormone receptor.
XX
PS Claim 2; Page; 123pp; English.

XX The present invention relates to a new method of treating an organism for
CC an inflammatory immune disease. The method of the invention comprises
CC reducing expression of a NURR subfamily nucleic acid sequence or
CC corticotropin releasing hormone (CRH) receptor nucleic acid sequence,
CC inhibiting transcriptional activity of a NURR superfamily member/CRH
CC receptor amino acid sequence, or reducing the level of NURR superfamily
CC member/CRH receptor sequence. The method is useful for treating an
CC organism for an inflammatory immune disease such as chronic inflammatory
CC joint disease, preferably arthritis, selected from rheumatoid arthritis,
CC psoriatic arthritis and sarcoid arthritis, ulcerative colitis and
CC thyroiditis. The method is also useful for screening a compound that
CC interferes with interaction of a NURR subfamily polypeptide with a
CC ligand, or identifying a compound for the treatment of an inflammatory
CC immune response. The agonist of the invention is useful for inhibiting
CC transcriptional activity of nuclear receptor polypeptide and the
CC antagonist is useful for decreasing the expression of a NURR subfamily
CC member. The present amino acid sequence represents the human nuclear
CC receptor NOR1 protein of the invention. This sequence is an antagonist of
CC the invention. Note: The present sequence is not shown in the
CC specification but was obtained from Genbank (7441771)
XX
SQ Sequence 625 AA;

Query Match 98.9%; Score 3299.5; DB 5; Length 625;
Best Local Similarity 99.2%; Pred. No. 6.4e-230;
Matches 621; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 MPCVQAQYSPSPGSSYAAQTSYSSYTTIIMNPDYTKLTMDLGSSTIATATTSLSIST 60
Db 1 MPCVQAQYSPSPGSSYAAQTSYSSYTTIIMNPDYTKLTMDLGSSTIATATTSLSIST 60

QY 61 FVEGYSSNYELKPCVYQWQRLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120
Db 61 FVEGYSSNYELKPCVYQWQRLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120

QY 121 SPDEVLSTSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180
Db 121 SPDEVLSTSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180

QY 181 PTVAGARFPLFHFKSPSPHPAPSPAGGHHGLGYDPTAAAGLSLPLGAAAAGSQAAL 240
Db 181 PTVAGARFPLFHFKSPSPHPAPSPAGGHHGLGYDPTAAAGLSLPLGAAAAGSQAAL 240

QY 241 HPYGLPLAKRAAPLAPFPPLGLTPSTASSLLGESPLSPSPSSSGEGTCAVGDNA 300
Db 241 HPYGLPLAKRAAPLAPFPPLGLTPSTASSLLGESPLSPSPSSSGEGTCAVGDNA 300

QY 301 COHYGVRTCEGCKGFFKRTVQKNKVCVLANKNCPVDKRRNRQYCRFQKCLSVGMWKE 360
Db 301 COHYGVRTCEGCKGFFKRTVQKNKVCVLANKNCPVDKRRNRQYCRFQKCLSVGMWKE 360

QY 361 VRTDSLKGRRLPSPKSPSQEPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420
Db 361 VRTDSLKGRRLPSPKSPSQEPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420

Db 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPSPICMNLVRLTDSPTPRDLDSRYC 420
 QY 421 PTDQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 480
 Db 421 PTDQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 479
 QY 481 RLSIRNTAEDKPVFCNGVLVHLRQLRGFGWLDSDIKDFSLNLSLNDIOALACLIAL 540
 Db 480 RLSIRNTAEDKPVFCNGVLVHLRQLRGFGWLDSDIKDFSLNLSLNDIOALACLIAL 539
 QY 541 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKTCTIGLOR 600
 Db 540 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKTCTIGLOR 599
 QY 601 IFYKLEDLVSPSPSIIDKFLDLPF 626
 Db 600 IFYKLEDLVSPSPSIIDKFLDLPF 625

RESULT 8

ADP28878
 ID ADF28878 standard; protein; 625 AA.
 AC ADF28878;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Human nor-1 nuclear receptor polypeptide 7441771.
 XX
 KW Human; nor-1; nuclear receptor; receptor; leukaemia; gene therapy;
 KW cytostatic; haematopoietic cell.
 XX
 OS Homo sapiens.
 XX
 PN WO2003088812-A2.
 XX
 PD 30-OCT-2003.
 XX
 PF 15-APR-2003; 2003WO-US011804.
 XX
 PR 17-APR-2002; 2002US-0373238P.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 PA (UNIW) UNIV WASHINGTON.
 XX
 PI Mullican SE, Conneely OM, Milbrandt J;
 XX
 XX WPI: 2003-854017/79.
 DR GENBANK; 7441771.
 DR

PT Inhibiting proliferation of a hematopoietic cell, useful for treating or
 PT preventing leukemia, comprises modulating the level or activity of nor-1
 PT and/or nur77 nuclear receptor.

PS Disclosure; SEQ ID NO 13; 101pp; English.

XX
 CC The present sequence is that of human nor-1 nuclear receptor polypeptide
 CC 7441771. Nor-1 has been identified as a molecular target for therapeutic
 CC intervention in the treatment of myeloid leukemia. The invention
 CC provides methods of inhibiting the proliferation of a haematopoietic stem
 CC cell or a haematopoietic myeloid cell by modulating the level of a nor-1
 CC and/or nur77 nuclear receptor. This involves increasing the level of the
 CC receptor polypeptide or polynucleotide, e.g. by administration of a
 CC vector comprising the polynucleotide. A claimed method of treating
 CC leukaemia comprises modulating a nor-1 and/or nur77 nuclear receptor in a
 CC haematopoietic stem cell or myeloid cell. Also claimed are methods of
 CC identifying an upregulator of expression of nor-1 and/or nur77, of
 CC identifying a compound for the treatment of leukaemia, and of screening
 CC for a compound for treatment of leukaemia, and a mouse model for
 CC leukaemia comprising a mouse having defective nor-1 and nur77 nucleic
 CC acid sequences.

XX Sequence 625 AA;

XX

Query Match 98.9%; Score 3299.5; DB 7; Length 625;
 Best Local Similarity 99.2%; Pred. No. 6.4e-220;
 Matches 621; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
 QY 1 MPCVQAQYSPSPSPGSSYAAQTYSSSEYTTIMNPDYTKLTMDLGSSTETATATSLPSIST 60
 Db 1 MPCVQAQYSPSPSPGSSYAAQTYSSSEYTTIMNPDYTKLTMDLGSSTETATATSLPSIST 60
 QY 61 FVEGYSSNYELKPCVYQMORELIKVGEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120
 Db 61 FVEGYSSNYELKPCVYQMORELIKVGEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120
 QY 121 SSPDEVLPSSTMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIATGCLLDPPMKAV 180
 Db 121 SSPDEVLPSSTMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIATGCLLDPPMKAV 180
 QY 181 PTVAGARFPLFHFKPSPPHPPAPSPAGGHHILGYDPTAAALSLPLGAAAAGSQAALLES 240
 Db 181 PTVAGARFPLFHFKPSPPHPPAPSPAGGHHILGYDPTAAALSLPLGAAAAGSQAALLES 240
 QY 241 HPYGLPLAKRAAPLAFPLGLTPTASLLGESPLSPSPSPSSSSSGEGTCAVCGDAA 300
 Db 241 HPYGLPLAKRAAPLAFPLGLTPTASLLGESPLSPSPSPSSSSSGEGTCAVCGDAA 300
 QY 301 COHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVVDKRRNRRCQYCRFKCLSVGMVKE 360
 Db 301 COHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVVDKRRNRRCQYCRFKCLSVGMVKE 360
 QY 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPSPICMNLVRLTDSPTPRDLDSRYC 420
 Db 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPSPICMNLVRLTDSPTPRDLDSRYC 420
 QY 421 PTDQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 480
 Db 421 PTDQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 479
 QY 481 RLSIRNTAEDKPVFCNGVLVHLRQLRGFGWLDSDIKDFSLNLSLNDIOALACLIAL 540
 Db 480 RLSIRNTAEDKPVFCNGVLVHLRQLRGFGWLDSDIKDFSLNLSLNDIOALACLIAL 539
 QY 541 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKTCTIGLOR 600
 Db 540 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKTCTIGLOR 599
 QY 601 IFYKLEDLVSPSPSIIDKFLDLPF 626
 Db 600 IFYKLEDLVSPSPSIIDKFLDLPF 625

RESULT 9

ADP28882
 ID ADF28882 standard; protein; 643 AA.
 AC ADF28882;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Pig nor-1 nuclear receptor polypeptide CAA09763.
 XX
 KW Pig; nor-1; nuclear receptor; receptor; leukaemia; gene therapy;
 KW cytostatic; haematopoietic cell.
 XX
 OS Sus scrofa.
 XX
 PN WO2003088812-A2.
 XX
 PD 30-OCT-2003.
 XX
 PF 15-APR-2003; 2003WO-US011804.
 XX
 PR 17-APR-2002; 2002US-0373238P.
 XX

(BAYU) BAYLOR COLLEGE MEDICINE.
(UNIW) UNIV WASHINGTON.

Mullican SE, Conneely OM, Milbrandt J;
WPI, 2003-854017/79.
GENBANK; CAA09763.

Inhibiting proliferation of a hematopoietic cell, useful for treating or preventing leukemia, comprises modulating the level or activity of nor-1 and/or nur77 nuclear receptor.

Disclosure; SEQ ID NO 17; 101pp; English.

The present sequence is that of pig nor-1 nuclear receptor polypeptide CAA09763. Nor-1 has been identified as a molecular target for therapeutic intervention in the treatment of myeloid leukemia. The invention provides methods of inhibiting the proliferation of a haematopoietic stem cell or a haematopoietic myeloid cell by modulating the level of a nor-1 and/or nur77 nuclear receptor. This involves increasing the level of the receptor polypeptide or polynucleotide, e.g. by administration of a vector comprising the polynucleotide. A claimed method of treating leukaemia comprises modulating a nor-1 and/or nur77 nuclear receptor in a haematopoietic stem cell or myeloid cell. Also claimed are methods of identifying an upregulator of expression of nor-1 and/or nur77, of identifying a compound for the treatment of leukaemia, and of screening for a compound for treatment of leukaemia, and a mouse model for leukaemia comprising a mouse having defective nor-1 and nur77 nucleic acid sequences.

Sequence 643 AA;

Query Match 93.9%; Score 3133.5; DB 7; Length 643;
Best Local Similarity 92.4%; Pred. No. 2.1e-208;
Matches 594; Conservative 10; Mismatches 22; Indels 17; Gaps 3;

QY 1 MPCVQAQYSPSPGSSYAAQYTSSEYTEINPDYTKLTWDLGSTTEITATATSLPST 60
Db 1 MPCVQAQYSPSPGSSYAAQYTSSEYTEINPDYTKLTWDLGSTTEITATATSLPST 60

QY 61 FVEGYSSNVELKPSVCYQMQ----RPLIKVEEGRAPSYHHHHHHHHHHHHHHHQ----- 110
Db 61 FMEGYSSNVELKPSCLYQMQSGRPRLIKMEEGRAHYHHHHHHHHHHHHHHHHHHHHHH 120

QY 111 --QHQQSTIPASSPEDEVLTSTNMYFKQSPSTPTTPAPPPQAGALWDEALPSAGGCTA 168
Db 121 PPOQQQSPSTPPSPGDEVLTSTNMYFKQSPSTPTTPFPQAGALWDEALPSAQGCTA 180

QY 169 PGPLDPPMKAVPTVAGARFPLFHFKSPHPHPAPSPAGGHILGYDPTAAALSLPLG-- 226
Db 181 PGPLDPPMKAVPTVAGARFPLFHFKSPHPHPAPSPAGGHILGYDPTAAALGLPLGAA 240

QY 227 ---AAAAAGSQAAALESHYPGLPLAKRAAPLAFPLGLTSPSTASSILGESLSLSPSPSR 283
Db 241 AAAAAAGSQAAALLEGHPYGLPLAKRAAALAFPLGLTTSPTSSLLGESLSLSPSPNR 300

QY 284 SSSSGEGTCVACGDNAACOHYGVRCCEGCKGFFKRTVQKNKXVCLANKKCPVDKRRNR 343
Db 301 STASGEGTCVACGDNAACOHYGVRCCEGCKGFFKRTVQKNKXVCLANKKCPVDKRRNR 360

QY 344 CQYCRQKCLSVGMVKVEVVRTDSLKGRGRPLFSKPSPLQOEPSQSPSPSPPICTMMNALV 403
Db 361 CQYCRQKCLSVGMVKVEVVRTDSLKGRGRPLFSKPSPLQOEPSQSPSPSPPICTMMNALV 420

QY 404 RALTDTSPRDLVSRYCPTDQAAGTDAHVQOQFNLLTASIDVSRWSAEKIPEFTDLPK 463
Db 421 RALTDTSPRDLVSRYCPTDQAAGTDAHVQOQFNLLTASIDVSRWSAEKIPEFTDLPK 480

QY 464 EDQTLTIESAFLELFLVRLSIRNSAEDKFCVFCNGLVHLRQLCRGFGEWLDISKDFSLN 523
Db 481 EDQTLTIESAFLELFLVRLSIRNSAEDKFCVFCNGLVHLRQLCRGFGEWLDISKDFSLR 540

QY 524 LOSLNLDIQALACLALSIMTERHGLKEPVRVEELCNKITSSLDKHOSKQALEPTESKV 583

Db	301	RRNRQYCRFOKCLSVGMVKEVWRTDSLKRRGRPLSKPKSPLOQEPSQSPSPPI	CWM	360
Qy	400	NALVRALTDSTPRDLIDYRYCPTDQAAAGTDAEHVQYFYNLLTASIDVSRSWAEKIPGFT		459
Db	361	NALVRALTDSTPRDLIDYRYCPTDQAAAGTDAEHVQYFYNLLTASIDVSRSWAEKIPGFT		420
Qy	460	DLPKEDQTLLESAPLEFLFVLRLSIRSNATADKFCVNCGLVHLRLOCLRGFGWLDISKD		519
Db	421	DLPKEDQTLLESAPLEFLFVLRLSIRSNATADKFCVNCGLVHLRLOCLRGFGWLDISKD		480
Qy	520	PSLNQSLNLDIQALACLSALSMITERHGLKEPKRVEELCNKITSSLDKHOSKQALEPT		579
Db	481	PSLNQSLNLDIQALACLSALSMITERHGLKEPKRVEELCNKITSSLDKHOSKQALEPT		540
Qy	580	ESKVLGALVELRKICTGLQRIFYKLKEDLVSPSPSIIDKLFDTLPF		626
Db	541	ESKVLGALVELRKICTGLQRIFYKLKEDLVSPSPSIIDKLFDTLPF		587
RESULT 12				
AAR92057				
ID	AAR92057 standard; protein; 628 AA.			
XX				
AC	AAR92057;			
XX				
DT	19-JUL-1996 (first entry)			
XX				
DE	Apoptotic cerebral neuron nuclear receptor protein.			
XX				
KW	Rat nuclear receptor gene; embryo; cerebral neuron; induction; apoptosis;			
KW	amplification; primer; PCR; Antisense; ribozyme; nerve; disease.			
XX				
OS	Rattus norvegicus.			
XX				
FH	Key Location/Qualifiers			
FT	Domain 293..628			
FT	/note= "DNA binding domain"			
XX				
PN	JF08023980-A.			
XX				
PD	30-JAN-1996.			
XX				
PF	15-JUL-1994; 94JP-00164434.			
XX				
PR	15-JUL-1994; 94JP-00164434.			
XX				
PA	(TERU) TERUMO CORP.			
XX				
XX	WPI; 1996-133421/14.			
DR	N-PSDB; AAT16151.			
XX				
PT	Nuclear receptor expressed in apoptosis of cerebral neuron - used to			
PT	design anti-sense oligo:nucleotide(s) and ribozyme(s) for treatment of			
PT	cerebral nerve disease etc.			
XX				
PS	Claim 1; Page 7-10; 11pp; Japanese.			
XX				
CC	This is the amino acid sequence of the novel rat nuclear receptor protein			
CC	designated NUR88. The gene was isolated from 16 day old rat embryo			
CC	cerebral neurons, induced into apoptosis, by amplifying the sequence from			
CC	cDNA using primers AAR28371-4. Antisense nucleic acids and ribozymes			
CC	targeted to the gene can be used as agents for treating cerebral nerve			
CC	diseases			
XX				
SQ	Sequence 628 AA;			
Query Match				
Best Local Similarity 92.1%; Score 3074; DB 2; Length 628;				
Matches 582; Conservative 12; Mismatches 29; Indels 8; Gaps 4;				
Qy	1 MPCVQAOYSPSPGSSVAAQYSSSEYTTTEINPNPYKLTMDLSTGTEITATATSLPISIT			
60				
Db	1 MPCVQAOYSPSPGSSVAAQYSSSEYTTTEINPNPYKLTMDLSTGTEITATATSLPISIT			
60				

Qy	61	FVEGYSSNVELKPSCVYQM---QRPLKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH	116
Db	61	FMEGYSSCELKPSCLYQMPSPGPRELKWEGREHYHHHHHHHHHHHHHHHHHHHHHHHH	118
Qy	117	IPPASSPEDEVLPSTSMYFKQSPPTTTPAPFPQAGALWDEALPSAPGCIAPGPLDDP	176
Db	119	IPPPSGPEDEVLPSTSMYFKQSPPTTTPGPPQAGALWDEALPSAPGCIAPGPLDDP	178
Qy	177	MKAVPTV-AGARPLPHEKPSPEHPAPSPAGGHHLYDPTAAALSLPLGAAAAAGSQA	235
Db	179	MKAVPPMAAARPFIF-FRPSPEHPAPSPAGGHHLYDPTAAALSLPLGAAAAAGSQA	237
Qy	236	AALESHPIGLPLAKRAAPLAPFPPLGLTPGTASSLLGESPSLSPSPSRSSSGEGTC	295
Db	238	AALSHPIGLPLAKRTATLTFFPLGLTASPTASSLLGESPSLSPSPSRSSSGEGTC	297
Qy	296	GDNAACQHYGVRTCEGCKGFFKTVQNAKYVCLANKNCVDRKRRNRCCYCFQKCLSV	355
Db	298	GDNAACQHYGVRTCEGCKGFFKTVQNAKYVCLANKNCVDRKRRNRCCYCFQKCLSV	357
Qy	356	GMVKEVVRTDSLKRRGRPLSKPKSPLOQEPSQSPSPPICMNNALVRALTDSTPRDLD	415
Db	358	GMVKEVVRTDSLKRRGRPLSKPKSPLOQEPSQSPSPPICMNNALVRALTDSTPRDLD	417
Qy	416	YSRYCPTDQAAAGTDAEHVQYFYNLLTASIDVSRSWAEKIPGFTDLPKEDQTLLES	475
Db	418	YSRYCPTDQATAGTDAEHVQYFYNLLTASIDVSRSWAEKIPGFTDLPKEDQTLLES	477
Qy	476	ELFVRLSIRSNATADKFCVNCGLVHLRLOCLRGFGWLDISKDFSLNLSLNDIQALA	535
Db	478	ELFVRLSIRSNATADKFCVNCGLVHLRLOCLRGFGWLDISKDFSLNLSLNDIQALA	537
Qy	536	CLSALSMITERHGLKEPKRVEELCNKITSSLDKHOSKQALEPTESKVLGALVELRK	595
Db	538	CLSALSMITERHGLKEPKRVEELCNKITSSLDKHOSKQALEPTESKVLGALVELRK	597
Qy	596	LGQRIFYKLKEDLVSPSPSIIDKLFDTLPF	626
Db	598	QGQRIFYKLKEDLVSPSPSIIDKLFDTLPF	628
RESULT 13			
ABB98438			
ID	ABB98438 standard; protein; 628 AA.		
XX			
AC	ABB98438;		
XX			
DT	21-OCT-2002 (first entry)		
XX			
DE	Murine Neural Orphan Receptor 1, NOR1, #2.		
XX			
KW	Murine; NOR1; neuroleptic; Neural Orphan Receptor 1;		
KW	Translocated in Extraskeletal Chondrosarcoma; TEC; cell proliferation;		
KW	apoptosis; central nervous system; transgenic mouse;		
KW	pain response threshold; impaired balance; impaired motor coordination.		
XX			
OS	Mus musculus.		
XX			
PN	WO200246391-A2.		
XX			
PD	13-JUN-2002.		
XX			
PF	05-DEC-2001; 2001WO-US047790.		
XX			
PR	06-DEC-2000; 2000US-0251794P.		
PR	24-SEP-2001; 2001US-0324614P.		
XX			
PA	(DELTA-) DELTAGEN INC.		
XX			
PI	Guenther C, Allen KD;		
XX			

Db 61 FMEGYSSCEKPSCLYQMPSSGRPLKMEEGREHYHHHHHHHHHHQO--QOPS 118
117 IPPASSPEDEVLPSTSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPP 176
119 IPPPSGPEDEVLPSTSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPP 178
177 MKAVPTV-AGARFPLFHFKPSPPHPPAPSPAGGHHGLGYDPTAAALSLPLGAAAAGSOA 235
179 MKAVPMAAAARFPFIF-FKPSPPHPPAPSPAGGHHGLGYDPTAAALSLPLGAAAAGSOA 237
236 AALESHPYGLPLAKRAAPLAPPLGLTPSTASSLLGESPSLPSPSSSSSGSGTCVAVC 295
238 AALEGHPYGLPLAKRTATLTTPPLGLTASPTASSLLGESPSLPSPSSSSSGSGTCVAVC 297
296 GDNAAQOHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDPKRRNRRCQYCRFQKCLSV 355
298 GDNAAQOHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDPKRRNRRCQYCRFQKCLSV 357
356 GMVKEVVRTDSLKGRGRPLSPKPSPLQOEPSQSPSPSPPICMNNALVRALTDSTPRDL 415
358 GMVKEVVRTDSLKGRGRPLSPKPSPLQOEPSQSPSPSPPICMNNALVRALTDSTPRDL 417
416 YSRYPCTDQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLESAPL 475
418 YSRYPCTDQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLESAPL 477
476 ELFVLRSLRSNTAEDKVFVFCNGLVLRLOCLRGFGWLDISKDFSLNQLSLNLDIOALA 535
478 ELFVLRSLRSNTAEDKVFVFCNGLVLRLOCLRGFGWLDISKDFSLNQLSLNLDIOALA 537
536 CLSALSMITERHGLKPEKVEELCNKITSSLDKHQKQALPEPTSKVLGALVELRKICT 595
538 CLSALSMITERHGLKPEKVEELCNKITSSLDKHQKQALPEPTSKVLGALVELRKICT 597
596 LGLQRIFFYLKLEDLVSPSPVIDKFLDLP 626
598 LGLQRIFFYLKLEDLVSPSPVIDKFLDLP 628

RESULT 15
ID ADF28880
AC ADF28880 standard; protein; 628 AA.
DT 12-FEB-2004 (first entry)
DE Rat nor-1 nuclear receptor polypeptide JC2493.
KW Rat; nor-1; nuclear receptor; receptor; leukaemia; gene therapy;
KW Cytostatic; haematopoietic cell.
XX Rattus sp.
XX WO2003088812-A2.
XX 30-OCT-2003.
XX 15-APR-2003; 2003WO-US011804.
XX 17-APR-2002; 2002US-0373238P.
XX (BAYU) BAYLOR COLLEGE MEDICINE.
XX (UNIW) UNIV WASHINGTON.
XX Mullican SE, Conneely OM, Milbrandt J;
XX WPI; 2003-854017/79.
XX Inhibiting proliferation of a hematopoietic cell, useful for treating or
PT preventing leukemia, comprises modulating the level or activity of nor-1
PT and/or nur77 nuclear receptor.

XX
PS
XX
CC The present sequence is that of rat nor-1 nuclear receptor polypeptide
CC JC2493. Nor-1 has been identified as a molecular target for therapeutic
CC intervention in the treatment of myeloid leukaemia. The invention
CC provides methods of inhibiting the proliferation of a haematopoietic stem
CC cell or a haematopoietic myeloid cell by modulating the level of a nor-1
CC and/or nur77 nuclear receptor. This involves increasing the level of the
CC receptor polypeptide or polynucleotide, e.g. by administration of a
CC vector comprising the polynucleotide. A claimed method of treating
CC leukaemia comprises modulating a nor-1 and/or nur77 nuclear receptor in a
CC haematopoietic stem cell or myeloid cell. Also claimed are methods of
CC identifying an upregulator of expression of nor-1 and/or nur77, of
CC identifying a compound for the treatment of leukaemia, and of screening
CC for a compound for treatment of leukaemia, and a mouse model for
CC leukaemia comprising a mouse having defective nor-1 and nur77 nucleic
CC acid sequences.
XX
SQ Sequence 628 AA;
Query Match 92.1%; Score 3074; DB 7; Length 628;
Best Local Similarity 92.2%; Pred. No. 2.7e-204;
Matches 582; Conservative 12; Mismatches 29; Indels 8; Gaps 4;
QY 1 MPCVQAQYSPSPSSVAAOTYSSEYTTIMNPDYKLTMDLSTETATATSLPSIST 60
DB 1 MPCVQAQYSPSPSSVAAOTYSSEYTTIMNPDYKLTMDLSTETATATSLPSIST 60
QY 61 FVEGYSNVELKPSCVYQM---ORPLIKVEEGRAPSYHHHHHHHHHHHHHHHHHHQO 116
DB 61 FMEGYSSCEKPSCLYQMPSSGRPLKMEEGREHYHHHHHHHHHHHHHHHHHHQO 118
QY 117 IPPASSPEDEVLPSTSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPP 176
DB 119 IPPPSGPEDEVLPSTSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPP 178
QY 177 MKAVPTV-AGARFPLFHFKPSPPHPPAPSPAGGHHGLGYDPTAAALSLPLGAAAAGSOA 235
DB 179 MKAVPMAAAARFPFIF-FKPSPPHPPAPSPAGGHHGLGYDPTAAALSLPLGAAAAGSOA 237
QY 236 AALESHPYGLPLAKRAAPLAPPLGLTPSTASSLLGESPSLPSPSSSSSGSGTCVAVC 295
DB 238 AALEGHPYGLPLAKRTATLTTPPLGLTASPTASSLLGESPSLPSPSSSSSGSGTCVAVC 297
QY 296 GDNAAQOHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDPKRRNRRCQYCRFQKCLSV 355
DB 298 GDNAAQOHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDPKRRNRRCQYCRFQKCLSV 357
QY 356 GMVKEVVRTDSLKGRGRPLSPKPSPLQOEPSQSPSPSPPICMNNALVRALTDSTPRDL 415
DB 358 GMVKEVVRTDSLKGRGRPLSPKPSPLQOEPSQSPSPSPPICMNNALVRALTDSTPRDL 417
QY 416 YSRYPCTDQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLESAPL 475
DB 418 YSRYPCTDQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLESAPL 477
QY 476 ELFVLRSLRSNTAEDKVFVFCNGLVLRLOCLRGFGWLDISKDFSLNQLSLNLDIOALA 535
DB 478 ELFVLRSLRSNTAEDKVFVFCNGLVLRLOCLRGFGWLDISKDFSLNQLSLNLDIOALA 537
QY 536 CLSALSMITERHGLKPEKVEELCNKITSSLDKHQKQALPEPTSKVLGALVELRKICT 595
DB 538 CLSALSMITERHGLKPEKVEELCNKITSSLDKHQKQALPEPTSKVLGALVELRKICT 597
QY 596 LGLQRIFFYLKLEDLVSPSPVIDKFLDLP 626
DB 598 LGLQRIFFYLKLEDLVSPSPVIDKFLDLP 628

Search completed: March 12, 2005, 17:48:10
Job time : 173 secs


```
Db 372 COHGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDPVKRRNRNRCQYCRFOKCLSVGMVKE 431
Qy 361 VVRTDSLKGRGRPLPSKPKSPLOQEPSQSPSPPICMNNALVRALTDSTPRDLDSRYC 420
Db 432 VVRTDSLKGRGRPLPSKPKSPLOQEPSQSPSPPICMNNALVRALTDSTPRDLDSRYC 491
Qy 421 PTQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 480
Db 492 PTQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 551
Qy 481 RLSIRSTAEDKVFVFCNGLVLRHQLCLRGFGEWLDSIKDPSLNLSQSLNLDIOALACLSAL 540
Db 552 RLSIRSTAEDKVFVFCNGLVLRHQLCLRGFGEWLDSIKDPSLNLSQSLNLDIOALACLSAL 611
Qy 541 SMITERHGLKEPKRVEELCNKITSSLDKHSQKQALEPTESKVLGALVELRKICTLGLQR 600
Db 612 SMITERHGLKEPKRVEELCNKITSSLDKHSQKQALEPTESKVLGALVELRKICTLGLQR 671
Qy 601 IFYKLEDLVSPPSIIDKFLDTLPF 626
Db 672 IFYKLEDLVSPPSIIDKFLDTLPF 697
```

RESULT 2

```
US-09-949-016-6776
; Sequence 6776, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6776
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6776
```

```
Query Match 99.4%; Score 3316; DB 4; Length 626;
Best Local Similarity 99.4%; Pred. No. 2.4e-247;
Matches 622; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLTMDLGSTETATATSLPSIST 60
Db 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLTMDLGSTETATATSLPSIST 60
Qy 61 FVEGYSNNYELKPSQVQWQRLIKVEEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHHH 120
Db 61 FVEGYSNNYELKPSQVQWQRLIKVEEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHHH 120
Qy 121 SSPDEVLPTSMYFKQSPSTTTTAPFPQAGALWDEALPSAPGCIAPGILLDPPMKAV 180
Db 121 SSPDEVLPTSMYFKQSPSTTTTAPFPQAGALWDEALPSAPGCIAPGILLDPPMKAV 180
Qy 181 PTVAGARFPLFHFKEPSPHPHPPSPAGGHLGVDPTAAALSLPLGAAAAGSQAALLES 240
Db 181 PTVAGARFPLFHFKEPSPHPHPPSPAGGHLGVDPTAAALSLPLGAAAAGSQAALLES 240
Qy 241 HPYGLPLAKRAAFLAPPLGLTSPASSLLGESPSLPPSPSSSSSGEGTCAVCGDNA 300
Db 241 HPYGLPLAKRAAFLAPPLGLTSPASSLLGESPSLPPSPSSSSSGEGTCAVCGDNA 300
```

```
Qy 301 COHGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDPVKRRNRNRCQYCRFOKCLSVGMVKE 360
Db 301 COHGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDPVKRRNRNRCQYCRFOKCLSVGMVKE 360
Qy 361 VVRTDSLKGRGRPLPSKPKSPLOQEPSQSPSPPICMNNALVRALTDSTPRDLDSRYC 420
Db 361 VVRTDSLKGRGRPLPSKPKSPLOQEPSQSPSPPICMNNALVRALTDSTPRDLDSRYC 420
Qy 421 PTQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 480
Db 421 PTQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 480
Qy 481 RLSIRSTAEDKVFVFCNGLVLRHQLCLRGFGEWLDSIKDPSLNLSQSLNLDIOALACLSAL 540
Db 481 RLSIRSTAEDKVFVFCNGLVLRHQLCLRGFGEWLDSIKDPSLNLSQSLNLDIOALACLSAL 540
Qy 541 SMITERHGLKEPKRVEELCNKITSSLDKHSQKQALEPTESKVLGALVELRKICTLGLQR 600
Db 541 SMITERHGLKEPKRVEELCNKITSSLDKHSQKQALEPTESKVLGALVELRKICTLGLQR 600
Qy 601 IFYKLEDLVSPPSIIDKFLDTLPF 626
Db 601 IFYKLEDLVSPPSIIDKFLDTLPF 626
RESULT 3
US-09-949-016-6699
; Sequence 6699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6699
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6699
```

```
Query Match 52.5%; Score 1751; DB 4; Length 598;
Best Local Similarity 54.5%; Pred. No. 1.1e-126;
Matches 363; Conservative 74; Mismatches 121; Indels 108; Gaps 18;
Qy 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLTMDLGSTETATATSLP 56
Db 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLTMDLGSTETATATSLP 58
Qy 57 SISTFVEGYSNNYELKPSQVQWQRLIKVEEGRAPSYHHHHHHHHHHHHHHHHHHHHHH 112
Db 59 SFSTFMDNVSTCYDVKPPCLYQWPLSGQSSIKVEDIQWNYQQHSH----- 105
Qy 113 QQPSTPPPASSDEVLPLSTSMYFKQSPSTTTTAPFPQAGALWDEALPSAPGCIAPG 171
Db 106 -----LPPQS-----EEMWPHSGSVYIKPSPPTTTTGFQVQHSIPMWD 148
Qy 172 LLDLPMKAVPT-----VAGARFPLFHFKEPSPHPH-----PSPAG 207
Db 149 LNFHFNQYVATTHMIEQRKTPVSRLSLFSFKQSPGPTPVSSQCMRFDGPLHVPMPPEPAG 208
Qy 208 GHHLGVDPTAAALSLPLGAAAAGSQAALLESHPYGLP-LAKRAAPLAPPLGLTSPPT 266
Db 209 SHHV-----VDQGTFAVFNPIRKPSMGFPGLQI-----CH 239
```



```

Db 153 GCYFGLSGVDLSSSGSYGSCAPSPSTFQPPQLSPWDGSGFHFSQTYEG 212
QY 154 -ALWDEALPSAPGCIAPGLDPPMKAVPTVAGARFPLFHFKPSPHPAPSPAGGHLG 212
Db 213 LRANTEQLPKASG-----PPQ-----PPAFSFSPTGPSPS----- 244
QY 213 YDPTAAALSL-PLGAAAAGSQAALSHPHYGLFLAKRAAPLAPPLPLGLTPSTASSLL 271
Db 245 ---LAQSPKLPFSQATHQLG-----EGESYSM-----PTAPFGL-----APTSPHLE 284
QY 272 G-----ESPSPSPSPSSSSG--EGTCVACGDNAACQHYGVRTCEGCKGFFKRTVQKNK 325
Db 285 GSGILDTP-VTSTKARSAGPAGSEGRCAVCGDNASCQHYGVRTCEGCKGFFKRTVQKNK 343
QY 326 YVCLANKCPVDRRRNRQYCRFOKCLSVGMVKEVVRTDSLKGRRLPSKPKSPLOQE 385
Db 344 YICLANKDCPDKRRNRQYCRFOKCLAVGMVKEVVRTDSLKGRRLPSKPKQP----- 399
QY 386 PSQSPSPSPICMMNALVREALTDSTP--RDLDYSVCPTDQAAAG-TDAEHVQQFYNLIT 442
Db 400 -----PDASPANLLTSLVRAHLDSGPGSTAKLDYSKFQELVLPHFGKEDAGDVQOQFYDLS 454
QY 443 ASIDVSRSAEKIPGTDLPKEDQTLILIESAFLEFLVLRSLSTRSNTAEKDFVFCNGLVLH 502
Db 455 GSLEVRKWAELKIPGAELSPADQDLLESFALEFLVLRSLSTRSNTAEKDFVFCNGLVLH 514
QY 503 RLQCLRGFEWLDISKDFNSLQSLMDIOALACLALSMTIRHCLKEPKEVEELCNKI 562
Db 515 RLQCARGFWDISILAFSRSLSLVLDVPAFACLSALVLIITDRHGLQEPKVEELQNKI 574
QY 563 TSLKDHQSGQALEPTE---SKVLGALVELRKICTLGLQIRIFYLKLEDLVSPSIIDK 618
Db 575 ASCLKEHVA-AVAGEFQAPASCLSLGLKLPKELTCTQGLQIRIFYLKLEDLVSPSIIDK 633
QY 619 LFLDTLPF 626
Db 634 IFMDTLPF 641

RESULT 6
US-09-277-078-2
; Sequence 2, Application US/09277078
; Patent No. 6312949
; GENERAL INFORMATION:
; APPLICANT: Sakurada, Kazuhiro
; APPLICANT: Palmer, Theo
; TITLE OF INVENTION: REGULATION OF TYROSINE HYDROXYLASE
; TITLE OF INVENTION: EXPRESSION
; FILE REFERENCE: 07251/031001
; CURRENT APPLICATION NUMBER: US/09/277,078
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-277-078-2

```

```

Query Match 19.8%; Score 659.5; DB 3; Length 300;
Best Local Similarity 41.6%; Pred. No. 6.2e-43;
Matches 151; Conservative 41; Mismatches 74; Indels 97; Gaps 13;

QY 1 MPCVQAQYSSPPGSSAAQTY-----SSEYTTETMNDYTKLMDLGSTTATATSLP 56
Db 1 MPCVQAQYSSPPGSSAAQTY-----SSEYTTETMNDYTKLMDLGSTTATATSLP 58
QY 57 SISTFVEGYSNYELKPSVCYQM-----QRPLIKVEGRAPSYHHHHHHHHHHHHH 112
Db 59 SFSTFMDNYSTGYDVKPCLQMPILSGQSSIKVEDIQMNYQHSH----- 105

```

```

QY 113 QQPSIIPASSPEDEVLP-STSMYFKOSPSTTPTTAPFPQAGALWDEALPSAPGCIAPGP 171
Db 106 ---LPPQS---BEMPHSGSVYTKFSSPPTPTGQFQVQHSFPMWDD-----PGS 148
QY 172 LLDPKMKAVPT-----VAGARFPLFHFKPSPHPPPA-----PSPAG 207
Db 149 LHMHPQNYVATTHIMEORKTPVSRSLSLFSFKQSPGPTFVSQCMRFDGDLHVPNMPPEAG 208
QY 208 GHILGYDPTAAALSLPLGAAAAGSQAALSHPHYGLP-LAKRAAPLAPPLPLGLTPSPT 266
Db 209 SHHV-----LDGQTFVAVPNPIRKPSMGPFLQI-----GH 239
QY 267 ASSLLGSSPLSPSPSSSSSGEGTCVACGDNAACQHYGVRTCEGCKGFFKRTVQKNK 326
Db 240 ASQLL--DTQVPSPPSRGSPSNEGLCAVCGDNAAACQHYGVRTCEGCKGFFKRTVQKNK 297
QY 327 VCL 329
Db 298 VCL 300

RESULT 7
US-07-952-800-2
; Sequence 2, Application US/07952800
; Patent No. 5403925
; GENERAL INFORMATION:
; APPLICANT: OZATO, KEIKO
; TITLE OF INVENTION: A NEW MEMBER OF THE NUCLEAR HORMONE
; TITLE OF INVENTION: RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND AND TOWNSEND
; STREET: 1 MARKET PLAZA, STEUANT TOWER, SUITE 2000
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/952,800
; FILING DATE: 19920928
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-21-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 533 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-952-800-2

```

```

Query Match 16.0%; Score 533.5; DB 1; Length 533;
Best Local Similarity 29.7%; Pred. No. 6.8e-33;
Matches 151; Conservative 70; Mismatches 189; Indels 99; Gaps 17;

QY 118 PPASSPEDEVLPSTSMYFKOSPSTTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPM 177
Db 83 PDSSSNP--LP-----QGVPPSPGPPPLPSTA-----PSLGGSGAPP-----PPM 124
QY 178 KAVPTVAGARFPLFHFKPSPHPAPSPAGGHLGYDPTAAALSLPLGAAAAGSQAAL 237
Db 125 P--PPPLGSPFPVVISSSMGSGLPPPPAPFPFGFVSSPQINSTVSLPGGSG----- 174

```

QY 238 LESHVPLGLAKRAAPLAPPLGLTPSPASSLLGESPSLPSPSSSSSGEGTCVAGCD 297
DB 175 -----PPEDVFP-----PVLGVRGLHCPPPPGGPGAGKRLCAICGD 210
QY 298 NAACQHYGVRTCEGCKGFFKRTVQNAKYVCLANRCPVDKRRNRQYCRFQKCLSVGM 357
DB 211 RSSGKHVGYVCEGCKGFFKRTIRKDLTVYSCRNDKXCTVDKQRNRQYCRVQKCLATGM 270
QY 358 VKEVVRTSLKRGRLSKPSPLQOBPSQSPSPSPICMNNALVRAL----- 406
DB 271 KREAQVEERQRC-----KQDQD-GGEGAGAPBEMPVDRILEAEVAEQKSDQGVGP 322
QY 407 -----TDSTPRDLDSRYCPTDQAAAGTDAEHVQOQFYNLLTASIDVSRSAEKIIFGFTDL 461
DB 323 GTGCGSGSPND-----PVTNICOAD-----KQLFTLV-----EWAKRIHFPSL 363
QY 462 PKEDQTLIESAFLEFVLRLSIRNTAEDKEFVFCNGVLVHLRQC-LRGFGWELDSI-KD 519
DB 364 PLDDQVILLRAGWNELLTASFHSRSDVDRDGILLATGLHVRHNSAHSAGVGAIFDRVLTE 423
QY 520 FSLNQLSLNLDIQALACLALSMT-ERHGLKEPKRVEELCNKITSSSLKDH-QSKQOALE 577
DB 424 LVSKNRDMRMDKTELGLRAITLFPNDAGLSNPSEVEVLREKVVASLETYCKQKYPEQQ 483
QY 578 PTESKVLGALVELRICKTGLQRIFFYLKL 606
DB 484 GRFALLRLPALRSIGLKCLEHLFFFKL 512

RESULT 8

US-09-949-016-8178
; Sequence 8178, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8178
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(448)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-949-016-8178

Query Match 15.9%; Score 530.5; DB 4; Length 448;
Best Local Similarity 34.5%; Pred. No. 9.3e-33;
Matches 152; Conservative 64; Mismatches 181; Indels 43; Gaps 14;

QY 203 PSPAGGHLGYDPTAAALSPL--IGAAAAGSQAALSHYPYGLPLAKRAAPLAPPLG 260
DB 14 PTFGGHLLNGY-PVPYPAFFPPMLGGLSPG-----ALTLQHLFVSGYS----- 59
QY 261 LTPSP-----TASSLJGE-SPSPSPSSSSSGEGTCVAGCNAACQHYGVRTCEGCKGF 315
DB 60 -TPSPATTETQSSSEIEIVSPSPSPPLPRI---YKPCFVCQDKSGYHYGSACGCKGF 116
QY 316 FKRTVQNAKYVCLANRCPVDKRRNRQYCRFQKCLSVGMVKEVVRTSLKGRRLP 375
DB 117 FRRSIQKNVYCHRDKNKCIINKVTNRQYCRQLQKCFEVGNSKESVRNDRK----- 169

QY 376 SKPKSPLOEBSQSPSPSPICMNNALVR-ALTDSTPRDLDSRYCP--TDQAAAGTDAE 432
DB 170 -KKKEVPPECSSESYTLTPEVGELIEKVRKAHOETFPALCQLGKTYTNNSSQORVSLDID 228
QY 433 HVQOQFYNLLTASIDVSRSAEKIIFGFTDLPKBDQTLTLLTESAFLEFVLRLSIRNTAEDK 492
DB 229 LWDKESELSKCIITVEFAKQLPGFTTLTIADQITLLKAACLDLILILICITRYTPEQDT 288
QY 493 FVFCNGVLVHLRQCUR-GFGEWLDISKDFSLNQLSLNLDIQALACLALSMT-ERHGLK 550
DB 289 MTFSDGLTLNRQTMENAGFPLTLVFAFANQLLPLEMDDAETGLLSAICLICGRDQDLE 348
QY 551 EPKRVEELCNKITSSSLKDHQSKQOALEP-TESKVLGALVELRICKTGLQRIFFYLKLEDL 609
DB 349 QPDRVDMLEQPELLEALKVYVRKRRPSRPHMFFMLMKITDLRSISAKGAERVITLQWEP 408
QY 610 VSPSPSIIDKLF-----LDLTL 624
DB 409 GSMPLLIQEMLENSRGLDTL 428

RESULT 9

US-08-592-383-2
; Sequence 2, Application US/08592383
; Patent No. 5830750
; GENERAL INFORMATION:
; APPLICANT: Teal, S. and S.J. Collins
; TITLE OF INVENTION: "Hematopoietic Cell Lines Bearing Altered Retinoic Acid Recep
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; STREET: 2800 Pacific First Centre, 1420 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette-5.25 inch, 1.2Mb storage
; COMPUTER: IBM PC/386 Compatible
; OPERATING SYSTEM: MS-DOS 4.01
; SOFTWARE: Word for Windows 5.01-t
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,383
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/099,242
; FILING DATE: July 28, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Brogerick, Thomas F.
; REGISTRATION NUMBER: 31,332
; REFERENCE/DOCKET NUMBER: FHCR-1-7190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 1-206-682-8100; 1-206-224-0709(direct)
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; DESCRIPTION: sequence encoded by SEQ. ID. NO. 1
US-08-592-383-2

Query Match 15.9%; Score 530.5; DB 2; Length 462;
Best Local Similarity 34.5%; Pred. No. 9.6e-33;
Matches 152; Conservative 64; Mismatches 181; Indels 43; Gaps 14;
QY 203 PSPAGGHLGYDPTAAALSPL--IGAAAAGSQAALSHYPYGLPLAKRAAPLAPPLG 260
DB 9 PTFGGHLLNGY-PVPYPAFFPPMLGGLSPG-----ALTLQHLFVSGYS----- 54

TELEPHONE: (212) 977-9550
TELEFAX: (212) 644-0525
TELEX: (212) 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-02320A-4

Query Match 15.9%; Score 530.5; DB 5; Length 462;
Best Local Similarity 34.5%; Pred. No. 9.6e-33;
Matches 152; Conservative 64; Mismatches 181; Indels 43; Gaps 14;

QY 203 PSPAGCHLGYDPTAAALSLP-IGAAAAAGSQAALSHYPYGLPLAKRAAPLAPPPILG 260
DB 9 PTPGGCHLNGY-PVPYPAFFPPMLGGLSPG-ALTLQQLPVSYS-----54

QY 261 LTPSP-----TASSLIGE-SPSLPSPSRSSSGEGTCAVGDNAACQHYGVRTCEGCKGF 315
DB 55 -TPSPATTIETOSSSEIIVSPSPPLPRI--YKPCFVCQDKSGYHYGVSAECGCKGF 111

QY 316 FKRTVQNAKYVCLANKNCVDRNRNRCQYCRFOKCLSVGMVKEVVRTSLKRRGRRLP 375
DB 112 FRSTQKNVYVYCHRDKNCKIINKVTRNRCQYRLQKCEVGVSKESVRNRNK-----164

QY 376 SKPKSPLOQEPSQSPSPPICMNALVR-ALTDSTPRDLVSRYP--TDQAAAGTDAE 432
DB 165 -KKEVPKPECSYVTLTPEVCELKVKRKAQETFPALCQLGKYVNNSSQRVSLDID 223

QY 433 HVQPFYNLLTASIDVSRSWAEKIRFTDLPKEDOTLLIESALELFLVRLSIRSNATDK 492
DB 224 LWDKFSLSKCIITKVTFAKQLPQFTLTITADQITLKAACLDILIRICTRYTPEQDT 283

QY 493 FVFCNGLVLRHQCLLR-GFGEMLDSIKDFSLMLQSLNLDIQALACLSALSMIT-ERHGLK 550
DB 284 MTFDGLTLNRQTMHNAFGPLTDLVFAFANQLLPLEWDDATGILSAICLIGRQDL 343

QY 551 EPRVVEELCNKTTSLKHQSKQALEP-TESKVLGALVELRKICTLGLQRIFYKLJEDL 609
DB 344 QPDRVDMLOEPLLEALKVYVRKRRSPRPHMFPKMLKMTDLRSISAKGAERVITLKMEIP 403

QY 610 VSPSPIDKLF-----LDTL 624
DB 404 GMPPLIQEMLNSRGLDTL 423

RESULT 12

US-08-764-870-7

Sequence 7, Application US/08764870

Patent No. 6236946

GENERAL INFORMATION:

APPLICANT: Scanlan, Thomas S

APPLICANT: Baxter, John D

APPLICANT: Fletterick, Robert J

APPLICANT: Wagner, Richard L

APPLICANT: Kushner, Peter J

APPLICANT: Apriletti, James W

APPLICANT: West, Brian

TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESS: Cooley Godward

STREET: Five Palo Alto Square, 3000 El Camino Real

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

RESULT 13

US-08-980-115-7

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,870
FILING DATE: 13-DEC-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,540
FILING DATE: 13-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,543
FILING DATE: 13-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,606
FILING DATE: 14-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Jackie N
REGISTRATION NUMBER: 35,966
REFERENCE/DOCKET NUMBER: UCAL-246/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)843-5000
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 525 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-764-870-7

Query Match 15.9%; Score 530.5; DB 3; Length 525;

Best Local Similarity 29.2%; Pred. No. 1.1e-32;

Matches 150; Conservative 68; Mismatches 187; Indels 109; Gaps 17;

QY 118 PPASSEDEVLPSTSMYFKQSPPTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPM 177

DB 75 PDSSSPNP--LP-----QGVPPSPGPPPLPSTA-----PTLGSGGAPP---PPM 116

QY 178 KAVTVAGARPPLPHFKSPSPHPAPSPAGHHGLGYDPTAAALSLPLGAAAAAGSQA 237

DB 117 P--PPPLGSPFVSSSMGSPGLPPPAPPGFSGPVSSPQINSTVSLPGGGSG-----166

QY 238 LESHYPGLPLAKRAAPLAPPLGLTPTSPASSLLGESPLSPSSSSSGECTCAVCGD 297

DB 167 -----PPEDVXP-----PVLGVRGLHCPPPGPGGAGKRLCAICGD 202

QY 298 NAACOHYGVRTCEGCKGFFKRTVQNAKYVCLANKNCVDRNRNRCQYCRFOKCLSVGM 357

DB 203 RSSGHYGVYCEGCKGFFKRTIRKDLTYSCEDNKDCTVDKQRNRCQYCRFOKCLATGM 262

QY 358 VKEVVRTDSLKGR-----RGRLPSPKSPLOQEPSQSPSPPICMNALVRALTD- 408

DB 263 KREAVOEERQGRKDKDGEGCAGGAPEE-----MPVDRILEAEELAVEQKSDQ 309

QY 409 -----STPRDLDSRYCFTDQAAAGTDAEHVQOQFYNNLLTASIDVSRSWAEKIP 456

DB 310 GVEGPGGTGGSGSSPND-----PVTNICQAAD-----KQLFTLV-----EWAKRIP 350

QY 457 GFTDLPKEDQTLLESAFLFVLRLSIRSNATAEDKPFVFCNGLVLRHQLC-LRGFCWEILD 515

DB 351 HFSSLPDLDVILLRAGWNELLIASFSHRSDIVRGGILLATGLVHRNSAHAGVGAIFD 410

QY 516 SI-KDFSLNQLSLNLDIQALACLSALSMIT-ERHGLKEPKRVVEELCNKTTSLKHQ- 572

DB 411 RVLTELKSKMRDMRDKTELGLCLRAIILFNPDAKGLSNPSEVEVLRKRYASILETYCKOK 470

QY 573 GQALEPTESKVLGALVELRKICTLGLQRIFYKL 606

DB 471 YPEOQGRFAKLLRLPALRSIGLKCIEHLFFPKL 504

; Sequence 7, Application US/08980115

; Patent No. 6266622

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Thomas S.

; APPLICANT: Baxter, John D.

; APPLICANT: Fletcher, Robert J.

; APPLICANT: Wagner, Richard L.

; APPLICANT: Kushner, Peter J.

; APPLICANT: Apriletti, James W.

; APPLICANT: West, Brian L.

; APPLICANT: Shiao, Andrew K.

; TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS

; FILE REFERENCE: UCAL-246/02US

; CURRENT APPLICATION NUMBER: US/08/980,115

; CURRENT FILING DATE: 1997-11-26

; EARLIER APPLICATION NUMBER: 08/764,870

; EARLIER FILING DATE: 1996-12-13

; EARLIER APPLICATION NUMBER: 60/008,606

; EARLIER FILING DATE: 1995-12-14

; EARLIER APPLICATION NUMBER: 60/008,543

; EARLIER FILING DATE: 1995-12-13

; EARLIER APPLICATION NUMBER: 60/008,540

; EARLIER FILING DATE: 1995-12-13

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 525

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (292)..(523)

; OTHER INFORMATION: minimal ligand binding domain

; US-08-980-115-7

Query Match 15.9%; Score 530.5; DB 3; Length 525;
Best Local Similarity 29.2%; Pred. No. 1.1e-32;
Matches 150; Conservative 68; Mismatches 187; Indels 109; Gaps 17;

QY 118 PPASSPEDEVLPSTSMYFKQSPPTPTTPAPPPQAGALWDEALPSAPGCIAPGGLDPPM 177

DB 75 PDSSSPNP--LP-----QGVPPSPPPGPPPLPSTA-----PTLGGSGAPP--BPM 116

QY 178 KAVPTVAGARPLFHFKEPSPHPAPSPAGGHHLGYDPTAAALSLPLGAAAAAGSQA 237

DB 117 P--PPPLGSPFPVSSSMGSGPLPPAPPGSGPVSSPQINSTVSLPGGSG----- 166

QY 238 LESHYPGLPLAKRAAPLAPPLGLTPTPTASLLGESPLSPSPSRSSSGEGTCVCGD 297

DB 167 -----PPEDVKP-----PVLGVRGLHCPPPGPGAGKRLCAICGD 202

QY 298 NAAOCHYGVRTCEGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRQYCRFOKCLSVGM 357

DB 203 RSSGKHGYVSGEGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRQYCRFOKCLSVGM 262

QY 358 KVEVVRTDLSKGR-----RGRLPSPKSPLOQEPSPSPSPPICMNALVRALTD- 408

DB 263 KREAVQERQKDKDGCAGGAPEE-----MPVDRILEAEALVEQKSDQ 309

QY 409 -----STPRDLDSRYCPTDQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIP 456

DB 310 GVEGPGGTGGSGSPND-----PVTNICQAAD-----KQLFTLV-----EWAKRIP 350

QY 457 GTFDLPKEDOTLLIESAFLEVLRLSRSNTAEDKFVFCNGLVLRHLCQ-LRGGEWLD 515

DB 351 HESSLPDQVLLRAGMELLIAFSHRSIDVRDGIILLATGLVHRNSAHSAGVAIFD 410

QY 516 ST-KDFSLNQLSDIQAALCLALSMT-ERHGLKEPKRVEELCNKITSLSKDH-QSK 572

DB 411 RVLTSLVSKQRDRMDKTELGLCLRAILLNPNPAKGLSNPSEVEVLRKVASLETYCKOK 470

QY 573 GOALPETESKVLGALVELRKICTLGLQRIFYKL 606

Db 471 YPEOQGRFAKILLRLPALRSIGLKLHLEHFFFKL 504

RESULT 14

US-08-216-592A-4

; Sequence 4, Application US/08216592A

; Patent No. 6635429

; GENERAL INFORMATION:

; APPLICANT: LEID, MARK

; APPLICANT: KASTNER, PHILIPPE

; APPLICANT: CHAMON, PIERRE

; TITLE OF INVENTION: NOVEL HETERODIMERIC STEROID RECEPTOR

; TITLE OF INVENTION: PROTEINS, GENES ENCODING SAME, AND USAGE THEREOF

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox

; STREET: 1100 New York Avenue NW Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/216,592A

; FILING DATE: 23-MAR-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/825,667

; FILING DATE: 24-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: GOLDSTEIN, JORGE A

; REGISTRATION NUMBER: 29,021

; REFERENCE/DOCKET NUMBER: 1383.0060002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 533 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-216-592A-4

Query Match 15.9%; Score 530.5; DB 4; Length 533;

Best Local Similarity 29.5%; Pred. No. 1.2e-32;

Matches 150; Conservative 71; Mismatches 189; Indels 99; Gaps 17;

QY 118 PPASSPEDEVLPSTSMYFKQSPPTPTTPAPPPQAGALWDEALPSAPGCIAPGGLDPPM 177

DB 83 PDSSSPNP--LP-----QGVPPSPPPGPPPLPSTA-----PTLGGSGAPP--BPM 124

QY 178 KAVPTVAGARPLFHFKEPSPHPAPSPAGGHHLGYDPTAAALSLPLGAAAAAGSQA 237

DB 125 P--PPPLGSPFPVSSSMGSGPLPPAPPGSGPVSSPQINSTVSLPGGSG----- 174

QY 238 LESHYPGLPLAKRAAPLAPPLGLTPTPTASLLGESPLSPSPSRSSSGEGTCVCGD 297

DB 175 -----PPEDVKP-----PVLGVRGLHCPPPGPGAGKRLCAICGD 210

QY 298 NAAOCHYGVRTCEGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRQYCRFOKCLSVGM 357

DB 211 RSSGKHGYVSGEGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRQYCRFOKCLSVGM 270

QY 358 KVEVVRTDLSKGRRLSPKSPLOQEPSPSPSPPICMNALVRAL----- 406

DB 271 KREAVQERQK-----KDKDG-DGEGAGGAPPEMPVDRILEAEALVEQKSDQVEGP 322

QY 407 -----TDSTPRDLDSRYCPTDQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDL 461

Db 323 GGTGGSGSPND-----PVNICOAD-----KQLFTLV-----EWAKRIHFSSSL 363
Qy 462 PKEDQTLIESAFLELFLVLRIRNTAEDKVFVNGVLVLRHLC- LRGFGWLDI-KD 519
Db 364 PLDDQVILLRACWNEILLASFHSRSDVRDGIILLATGLHVRHNSAHSAGVGAIFDRVLTE 423
Qy 520 FSLNLSQSLNDIQALACLALSMT- ERHGLKEPKRVEELCNKITSSILKDH-QSKGQALE 577
Db 424 LVSKRDRMDKTELGCLRAILFNPDAKGLSNPSEVEVLRKRYASLETYCKQKYPBOQ 483
Qy 578 PTESKVLGALVELRKICTLGLQRIFYKL 606
Db 484 GRFAXLLRLPALRSIGLKLCHLFFFKL 512

RESULT 15
US-09-949-016-11572
; Sequence 11572, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11572
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11572

Query Match 15.9%; Score 530.5; DB 4; Length 577;
Best Local Similarity 29.4%; Pred. No. 1.3e-32;
Matches 151; Conservative 70; Mismatches 189; Indels 103; Gaps 17;
Qy 118 PPASSPEDEVLPSTMYFKQSPSTPTTTPAFPQAGALWDEALPSAPGCIAPGLDPPM 177
Db 123 PDSSSFNP--LP-----QGVPPSPPGPLPSTA-----PSLGGSGAPP--PPM 164
Qy 178 KAVTVAGARFPLFHEKSPHPHPAPSPAGHHLGYDPTAAALSLPLGAAAAAGSQA 237
Db 165 P--PPPLGSPFVVISSSMGSGLPPPPAPPGFSGVPSPQINSTVSLPGGSG-- 214
Qy 238 LESHVYGLPLAKRAAPLAPPPLGLTPSPATSSLSGESPSLPPSPSSSGGTCAVCGD 297
Db 215 -----PPEDVXP-----PVLGVRLHCPPPPGGCGAKRLCAICGD 250
Qy 298 NAAOHYGVRTCEGKGFRTVQKNAKYVCLANKCPVDKRRNRQYCRFQKLSVGM 357
Db 251 RSSGXHYGVYCEGKGFRTIRKDLTYSRDNKDCVTVDKQRNRCQYCRYQKCLATGM 310
Qy 358 VKEVVRTSLKRRRLSKPKSPLOQSPSQSPSPPICMNVALVRAL----- 406
Db 311 KREAVQEEQRQC-----KDKDG-DGEGAGGAPEEMPVDRILEAEAVEQKSDQGVGP 362
Qy 407 -----TDSFPRDLDSRYCPTQAAAGTDAEHVQQFYNLLTASIDVSRSWAEKIPGFTDL 461
Db 363 GGTGGSGSPND-----PVNICOAD-----KQLFTLV-----EWAKRIHFSSSL 403
Qy 462 PKEDQTLIESAFLELFLVLRIRNTAEDKVFVNGVLVLRHLC- LRGFGWLDI-KD 516
Db 404 PLDDQVILLRACWNEILLASFHSRSDVRDGIILLATGLHVRHNSAHSAGVGAIFDRSLR 463

Qy 517 -IKDFSLNLSQSLNDIQALACLALSMT- ERHGLKEPKRVEELCNKITSSILKDH-QSKG 573
Db 464 VTELVSQRDRMDKTELGCLRAILFNPDAKGLSNPSEVEVLRKRYASLETYCKQKY 523
Qy 574 QALEPTESKVLGALVELRKICTLGLQRIFYKL 606
Db 524 PEQOGRFAXLLRLPALRSIGLKLCHLFFFKL 556
Search completed: March 12, 2005, 17:52:49
Job time : 45 secs

THIS PAGE BLANK (USPTO)

Result No.	Query			Description	
	Score	Match	Length	ID	Description
1	3333	99.9	626	9	US-09-853-386-64
2	3333	99.9	626	15	US-10-414-080-14
3	3316	99.4	626	9	US-09-853-386-65
4	3316	99.4	626	9	US-09-853-386-96
5	3316	99.4	626	17	US-10-659-004-116
6	3299.5	98.9	625	9	US-09-853-386-63
7	3299.5	98.9	625	15	US-10-414-080-13
8	3133.5	93.9	643	9	US-09-853-386-70
9	3133.5	93.9	643	15	US-10-414-080-17
10	3123	93.6	587	15	US-10-755-889-64
11	3074	92.1	628	9	US-09-853-386-66
12	3074	92.1	628	9	US-09-853-386-68
13	3074	92.1	628	9	US-09-853-386-73
					Sequence 64, Appl
					Sequence 14, Appl
					Sequence 65, Appl
					Sequence 96, Appl
					Sequence 116, Appl
					Sequence 63, Appl
					Sequence 13, Appl
					Sequence 70, Appl
					Sequence 17, Appl
					Sequence 64, Appl
					Sequence 68, Appl
					Sequence 73, Appl

US-10-659-004-116

```
Query Match          99.4%; Score 3316; DB 17; Length 626;
Best Local Similarity 99.4%; Pred. No. 5e-216;
Matches 622; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLMDLGSSEITATATTSLSIST 60
DB 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLMDLGSSEITATATTSLSIST 60

QY 61 FVEGYSSNYELKPCSVYQWQRLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120
DB 61 FVEGYSSNYELKPCSVYQWQRLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120

QY 121 SSPDEVLPSTSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGLLDPKAV 180
DB 121 SSPDEVLPSTSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGLLDPKAV 180

QY 181 PTVAGARPLFHFKSPSPHPAPSPAGGHHLYDPTAAALSLPLGAAAAAGSQAALLES 240
DB 181 PTVAGARPLFHFKSPSPHPAPSPAGGHHLYDPTAAALSLPLGAAAAAGSQAALLES 240

QY 241 HPYGLPLAKRAAPLAPPLGLTPSTASSLLGESPSLSPSPSSSSGEGTCAVCDNAA 300
DB 241 HPYGLPLAKRAAPLAPPLGLTPSTASSLLGESPSLSPSPSSSSGEGTCAVCDNAA 300

QY 301 COHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDRKRRNRRCQYCRFKCLSVGMVKE 360
DB 301 COHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDRKRRNRRCQYCRFKCLSVGMVKE 360

QY 361 VVRTSLKGRGRPLSPKPSPLQBPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420
DB 361 VVRTSLKGRGRPLSPKPSPLQBPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420

QY 421 PTDOAAAGTDAHVQOOFYNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 480
DB 421 PTDOAAAGTDAHVQOOFYNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 480

QY 481 RLSIRNTAEDKVFVNCGLVLRQLCLRGFGWLDLSIKDLSLNQSLNLDIQALACLAL 540
DB 481 RLSIRNTAEDKVFVNCGLVLRQLCLRGFGWLDLSIKDLSLNQSLNLDIQALACLAL 540

QY 541 SMITRHGLKEPKRVEELCNKITSSLKHQSKGQALEPTESKVLGALVELRKICTLGLQR 600
DB 541 SMITRHGLKEPKRVEELCNKITSSLKHQSKGQALEPTESKVLGALVELRKICTLGLQR 600

QY 601 IFYKLEDLVSPSPSIIDKFLDLP 626
DB 601 IFYKLEDLVSPSPSIIDKFLDLP 626
```

RESULT 6

```
US-09-853-386-63
; Sequence 63, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Bresnahan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
; FILE REFERENCE: P01972US1
; CURRENT APPLICATION NUMBER: US/09/853,386
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/203645
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 625
; TYPE: PRT
; ORGANISM: HUMAN
```

US-09-853-386-63

```
Query Match          98.9%; Score 3299.5; DB 9; Length 625;
Best Local Similarity 99.2%; Pred. No. 6.5e-215;
Matches 621; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLMDLGSSEITATATTSLSIST 60
DB 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLMDLGSSEITATATTSLSIST 60

QY 61 FVEGYSSNYELKPCSVYQWQRLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120
DB 61 FVEGYSSNYELKPCSVYQWQRLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120

QY 121 SSPDEVLPSTSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGLLDPKAV 180
DB 121 SSPDEVLPSTSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGLLDPKAV 180

QY 181 PTVAGARPLFHFKSPSPHPAPSPAGGHHLYDPTAAALSLPLGAAAAAGSQAALLES 240
DB 181 PTVAGARPLFHFKSPSPHPAPSPAGGHHLYDPTAAALSLPLGAAAAAGSQAALLES 240

QY 241 HPYGLPLAKRAAPLAPPLGLTPSTASSLLGESPSLSPSPSSSSGEGTCAVCDNAA 300
DB 241 HPYGLPLAKRAAPLAPPLGLTPSTASSLLGESPSLSPSPSSSSGEGTCAVCDNAA 300

QY 301 COHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDRKRRNRRCQYCRFKCLSVGMVKE 360
DB 301 COHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDRKRRNRRCQYCRFKCLSVGMVKE 360

QY 361 VVRTSLKGRGRPLSPKPSPLQBPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420
DB 361 VVRTSLKGRGRPLSPKPSPLQBPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420

QY 421 PTDOAAAGTDAHVQOOFYNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 480
DB 421 PTDOAAAGTDAHVQOOFYNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFVL 480

QY 481 RLSIRNTAEDKVFVNCGLVLRQLCLRGFGWLDLSIKDLSLNQSLNLDIQALACLAL 540
DB 481 RLSIRNTAEDKVFVNCGLVLRQLCLRGFGWLDLSIKDLSLNQSLNLDIQALACLAL 540

QY 541 SMITRHGLKEPKRVEELCNKITSSLKHQSKGQALEPTESKVLGALVELRKICTLGLQR 600
DB 541 SMITRHGLKEPKRVEELCNKITSSLKHQSKGQALEPTESKVLGALVELRKICTLGLQR 600

QY 601 IFYKLEDLVSPSPSIIDKFLDLP 626
DB 601 IFYKLEDLVSPSPSIIDKFLDLP 626
```

RESULT 7

```
US-10-414-080-13
; Sequence 13, Application US/10414080
; Publication No. US2003020288A1
; GENERAL INFORMATION:
; APPLICANT: MULLICAN, SHANNON E.
; APPLICANT: CONNEELY, ORLA M.
; APPLICANT: MILBRANDT, JEFFREY
; TITLE OF INVENTION: NOR-1 AND NUR77 NUCLEAR RECEPTORS AS TARGETS FOR
; FILE REFERENCE: P02454US1
; CURRENT APPLICATION NUMBER: US/10/414,080
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/373,238
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-414-080-13
```


THIS PAGE BLANK (USPTO)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	3299.5	98.9	625	2	S71930	neuron-derived rec
2	3074	92.1	628	2	JC2493	neuron-derived orp
3	2032	60.9	430	2	S66671	neuron-derived rec
4	1751	52.5	598	2	A57040	T-cell nuclear rec
5	1746	52.3	598	2	A46225	nuclear orphan rec
6	1728	51.8	598	2	I84692	nuclear orphan rec
7	1367	41.0	598	2	A37251	probable nuclear h
8	1350.5	40.5	577	2	S33763	hormone receptor N
9	1316.5	39.5	601	1	QJWMSN1	probable hormone r
10	1277	38.3	563	2	QJQ623	nerve growth facto
11	1167	35.0	551	2	S58238	DHR38 protein - fir
12	1023	30.7	360	2	S58205	DHR38 protein - si
13	554.5	16.6	619	1	T43348	nuclear steroid ho
14	554	16.6	562	2	G88408	protein nhr-6 [imp
15	543	16.3	455	2	S13512	retinoic acid rece
16	531.5	15.9	462	2	S05050	retinoic acid nucl
17	530.5	15.9	462	1	A29491	retinoic acid rece
18	529.5	15.9	533	2	S37781	retinoid X recepto
19	525.5	15.7	446	2	A34418	H-2 region II bind
20	521.5	15.6	448	2	D41727	retinoid X recepto
21	521.5	15.6	520	2	I84718	RXR-beta1 isoform
22	521	15.6	955	4	C40045	probable transcrip
23	520.5	15.6	458	2	S06123	retinoic acid rece
24	519.5	15.6	451	2	A41651	retinoic acid rece
25	514.5	15.4	459	2	A41977	retinoic acid rece
26	514.5	15.4	476	2	B43977	retinoic acid rece
27	511.5	15.3	448	2	A43786	retinoic acid rece
28	509	15.3	441	2	I50515	retinoid X recepto
29	507	15.2	448	2	S05051	retinoic acid rece

	Query Match	98.9%; Score 3299.5; DB 2; Length 625;	
	Best Local Similarity 99.2%; Pred. No. 4.8e-174;		
	Matches 621; Conservative 1; Mismatches 3; Indels 1; Gaps 1;		
Qy	1 MPCVQAQYSPSPGSSYAAQTYSSEYTTEIMNPDYTKLTMDLGSTEITATATLSLPSIST 60 		
Dd	1 MPCVQAQYSPSPGSSYAAQTYSSEYTTEIMNPDYTKLTMDLGSTEITATATLSLPSIST 60 		
Qy	61 FVEGYSSNYELKSCVCYQMORPLIKVEEGRAPSYHHHHHHHHHHHQQOHOQPSP1PPA 120 		
Dd	61 FVEGYSSNYELKSCVCYQMORPLIKVEEGRAPSYHHHHHHHHHHHQQOHOQPSP1PPA 120 		
Qy	121 SSPDEVLPSTSMYFKQSPSPTTTTAPFPQAGALWDEALPSPAGC1APGLDPPMKAV 180 		
Dd	121 SSPDEVLPSTSMYFKQSPSPTTTTAPFPQAGALWDEALPSPAGC1APGLDPPMKAV 180 		
Qy	181 PTVAGARFPLFHFKPGPPHPAPSPAGCGHILGVDPDTAAAALSPLGAAAAAGSQAAALE 240 		
Dd	181 PTVAGARFPLFHFKPGPPHPAPSPAGCGHILGVDPDTAAAALSPLGAAAAAGSQAAALE 240 		
Qy	241 HPYGLPLAKRAAAPLAFFPLGLTPSPATSSLGGSPSLPSPRSSSSSGEGTCACVCGNAA 300 		
Dd	241 HPYGLPLAKRAAAPLAFFPLGLTPSPATSSLGGSPSLPSPRSSSSSGEGTCACVCGNAA 300 		
Qy	301 COHYGVRTCEGCKGFRTYQKNAYCYCLANKQCPCVDKKRRNRRCQYCRFOKLCSVGVMYKE 360 		

QY 117 IPPASSPEDEVLPSTSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGPIIDPP 176
|||
Db 119 IPPSGPEDEVLPSTSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGPIIDPPQ 178
|||
QY 177 MKAVPTV-AGARPLPLFHFKPSPPHPPAPSPAGGHHLYGYDPTAAALSLPLGAAAAAGSQA 235
|||
Db 179 MKAVPPMAAARFPFIF-FKPSPPHPPAPSPAGGHHLYGYDPTAAALSLPLGAAAAAGSQA 237
|||
QY 236 AALESHPYGLPLAKRAAPLAPPLGLTPSPPTASSLLGESPSLPSPPSRSSSGEGTCAVC 295
|||
Db 238 AALEGHPYGLPLAKRTATLTTPPLGLTASPTASSLLGESPSLPSPPNRRSSSGEGTCAVC 297
|||
QY 296 GDNAAACOHYGVRTCECKGFFKRTVOKNAKYVCLANKNCVPDKRRNRQCYCRFOKCLSV 355
|||
Db 298 GDNAAACOHYGVRTCECKGFFKRTVOKNAKYVCLANKNCVPDKRRNRQCYCRFOKCLSV 357
|||
QY 356 GMVKEVVRTDSLKGRRGLRPSKPSLQOEPSQSPSPSPPICMNALVRALTDSTPRDLD 415
|||
Db 358 GMVKEVVRTDSLKGRRGLRPSKPSLQOEPSQSPSPSPPICMNALVRALTDSTPRDLD 417
|||
QY 416 YSR 418
|||
Db 418 YSR 420
|||

RESULT 4

A57040
T-cell nuclear receptor NOT - human
C:Species: Homo sapiens (man)
C:Date: 03-Oct-1995 #sequence revision 03-Oct-1995 #text_change 09-Jul-2004
C:Accession: A57040; I56260; S38636
R:Magas, H.W.; Rilke, O.; Bravo, R.; Senger, G.; Kroczeck, R.A.
Mol. Endocrinol. 8, 1583-1591, 1994
A:Title: NOT, a human immediate-early response gene closely related to the steroid/thyroid
A:Reference number: A57040; MUID:95183071; PMID:7877627
A:Accession: A57040
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-598 <NAG>
A:Cross-references: UNIPROT:P43354; EMBL:X75918; NID:G415822; PIDN:CAA53518.1; PID:G4158
R:Okabe, T.; Takayanagi, R.; Imasaki, K.; Hajji, M.; Nawata, H.; Watanabe, T.
J. Immunol. 154, 3871-3879, 1995
A:Title: cDNA cloning of a NGFI-B/nur77-related transcription factor from an apoptotic h
A:Reference number: I56260; MUID:95221892; PMID:7706727
A:Accession: I56260
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 64-464, 'W', 466-467, 'W', 469-598 <RES>
A:Cross-references: GB:S7154; NID:G913966; PIDN:AAB33999.1; PID:G913967
C:Genetics:
A:Gene: GDB:NURR1; TINUR; NOT
A:Cross-references: GDB:624567; OMIM:601828
A:Map position: 2q22-2q23
C:Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog
C:Keywords: DNA binding; zinc finger
F:261-514/Domain: erba transforming protein homology <ERBA>

Query Match 52.5%; Score 1751; DB 2; Length 598;
Best Local Similarity 54.5%; Pred. No. 4.7e-89;
Matches 363; Conservative 74; Mismatches 121; Indels 108; Gaps 18;
QY 1 MPCVQAQYSPSPGSSYAAQTY-----SSEYTTIMNPDYTKLTMGLSTGTEITATATSLP 56
|||
Db 1 MPCVQAQYSGSPGASPASQSYSHSGEYSSDFLTPEFVKFSMDLTNTETI--TATTSLP 58
|||
QY 57 S1STFVEGVSSNYELKPSVCYQM---QRPLKVEGRAPSYHHHHHHHHHHHHHHHHHHQOQH 112
|||
Db 59 SFSTFMDNTSTGYDVKPCLYQMPLSGQSSIKVEDIQMNTYQSHS----- 105
|||
QY 113 QOFSIPPASSPEDEVLP-STSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGP 171
|||
Db 106 ----LPPQS---EEMPHSGSVYKPSSTPTTTCFQVQHSFPMWDD-----PGS 148
|||

QY 172 LLDPPMKAVPT-----VAGARPLPLFHFKSPSPHPPA-----PSPAG 207
|||
Db 149 LNFHQNYYATTHMIEQRTKTPVSRSLFSFKOSPPGTPVSSQCMRFDGLPHVPMNPEPAG 208
|||
QY 208 GHHLGYDPTAAALSLPLGAAAAAGSQAALSHSPYGLP-LAKRAAPLAPPLGLTPSP 266
|||
Db 209 SHHV-----VDQITFAVNPPIRKPASMGFPGLQI---GH 239
|||
QY 267 ASLLGESPSLPSPPSRSSSGEGTCAVCGDNAACOHYGVRTCECKGFFKRTVQKNKY 326
|||
Db 240 ASQLL--DTQVPSPPSRGSPSNEGLCAVCGDNAAACOHYGVRTCECKGFFKRTVQKNKY 297
|||
QY 327 VCLANKNCVPDKRRNRQCYCRFOKCLSVGMVKEVVRTDSLKGRRGLRPSKPSLQOEP 386
|||
Db 298 VCLANKNCVPDKRRNRQCYCRFOKCLAVGMVKEVVRTDSLKGRRGLRPSKPS-----P 352
|||
QY 387 SOPSPPSPPICMNALVRALTDSTP--ROLDYSRY-CPTDQAAAAGTDAEHVQFYQVNLTA 443
|||
Db 353 QEPSPPSPVSLISALVRAHVDSNPAMTSLDYSRQANPDYQMSGDDTQHIQQFYDLTLC 412
|||
QY 444 SIDVRSWAEEKIPGFTDLPKEDQTLIESAFLEFVLRSLRSINSNTAEDKPFVFCNGLVLHR 503
|||
Db 413 SMEIIRGWAEEKIPGFADLPKADQDLLFESAFLEFVLRSLAYRSNPVEGKLIFCNGVVLHR 472
|||
QY 504 LOCLRGCEWLDISKDFSLNLOSLNLDIOALCLSALSMITRHHGLKEPKRVEELCNKIT 563
|||
Db 473 LQCVRGFEWIDSIVEFSSNLQNMNIDISAFSICIAALAMVTERRHGLKEPKRVEELQNKIV 532
|||
QY 564 SSLKDHQ--SKQQALETPE-SKVLGALVELRIKICTLGLQRIFYKLKLELVSPSPSIDKLF 620
|||
Db 533 NCLKDHVTFNNGLNRPVLSKLLGLPELRLCTQGLQRIFYKLKLELVSPSPAIIDKLF 592
|||
QY 621 LDTLPP 626
|||
Db 593 LDTLPP 598
|||

RESULT 5

A46225
nuclear orphan receptor Nurrl - mouse
N:Alternate names: Nur-related transcription factor 1
C:Species: Mus musculus (house mouse)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A46225
R:Law, S.W.; Conneely, O.M.; DeMayo, F.J.; O'Malley, B.W.
Mol. Endocrinol. 6, 2129-2135, 1992
A:Title: Identification of a new brain-specific transcription factor, NURR1.
A:Reference number: A46225; MUID:93149122; PMID:1491694
A:Accession: A46225
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-598 <LAW>
A:Cross-references: UNIPROT:Q06219; GB:S53744; NID:G264410; PIDN:AAB25138.1; PID:G26441
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIN:123734, NCBI:123736)
C:Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog
C:Keywords: DNA binding; nucleus; transcription regulation; zinc finger
F:261-514/Domain: erba transforming protein homology <ERBA>

Query Match 52.3%; Score 1746; DB 2; Length 598;
Best Local Similarity 54.4%; Pred. No. 8.8e-89;
Matches 362; Conservative 75; Mismatches 121; Indels 108; Gaps 18;
QY 1 MPCVQAQYSPSPGSSYAAQTY-----SSEYTTIMNPDYTKLTMGLSTGTEITATATSLP 56
|||
Db 1 MPCVQAQYSGSPGASPASQSYSHSGEYSSDFLTPEFVKFSMDLTNTETI--TATTSLP 58
|||
QY 57 S1STFVEGVSSNYELKPSVCYQM---QRPLKVEGRAPSYHHHHHHHHHHHHHHHHHHQOQH 112
|||
Db 59 SFSTFMDNTSTGYDVKPCLYQMPLSGQSSIKVEDIQMNTYQSHS----- 105
|||
QY 113 QOFSIPPASSPEDEVLP-STSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGP 171
|||

Db 106 ----LPPQS---EEMPHSGSVYKPPSPPTSTPSTFQVQHSPWDD-----PGS 148

Qy 172 LLDPPMKAVPT-----VAGARPLPHFKSPHPPEA-----PSPAG 207

Db 149 LHNPHQNVATTHMIEQKTPVSRSLSPFSKQSPGTPGVSSCOMRFDGFLHVPNMPPEAG 208

Qy 208 GHHLGYDPTAAALSLPLGAAAAAGSQAALSHPHYGLP-LAKRAAPLAFPLGLTSPPT 266

Db 209 SHHV-----VDGQTFVAVNPPIRKPASMGFFGLQI---GH 239

Qy 267 ASSLLGESPSLPPSRSSSGEGTCVAGCDNAACQHYGVRTCEGCKGFFKRTVQKNKY 326

Db 240 ASQLL--DTQVPSPPSRGSPSNEGLCAVCGDNAACQHYGVRTCEGCKGFFKRTVQKNKY 297

Qy 327 VCLANKNCVDKRRNRRCQYCFQKCLSVGMVKEVVRTDSLKGRGRPLSPKPSLQOEP 386

Db 298 VCLANKNCVDKRRNRRCQYCFQKCLAVGMVKEVVRTDSLKGRGRPLSPKPS-----P 352

Qy 387 SOPSPPSPPICMNALVRALTDSTP--RDLDSRY--CPTDQAAAGTDAHVQOQFYNLLTA 443

Db 353 QDPSPPSPVSLISALVRAHVDSNPAMTSLDSYRFOANPDYQMSGDDTQHIOQFYDLLTG 412

Qy 444 SIDVRSWAEEKIPGFTDLPKEDQTLIESAFLEFVLRLSIRSNATDAEDKFFVFCNGLVLR 503

Db 413 SMEIIRGWAEEKIPGFADLPKADQDLIFESAFLEFVLRLAYRSNPVEGKLIFFCNGVLR 472

Qy 504 LQCLRGFGEWLSIDKDFSLNLSQSLNLDIOALACLSALSMTIRHGLKPKRVEELCNKIT 563

Db 473 LQCVRGFGEWLSIDKDFSLNLSQSLNLDIOALACLSALSMTIRHGLKPKRVEELCNKIT 532

Qy 564 SSKLDHQ--SKGOALEPTE--SKVIGALVELRKTCTGLQRIFFYLKLEDLVSPPSIIDKLF 620

Db 533 NCLKDHVTFNNGLNRPNYLSKLLGKLPRLTCTCQGLQRIFFYLKLEDLVPPPAIIDKLF 592

Qy 621 LDTLPF 626

Db 593 LDTLPF 598

RESULT 6.

184692

nuclear orphan receptor HZF-3 - rat

N;Alternate names: thyroid/steroid receptor homolog RNR-1

C;Species: Rattus norvegicus (Norway rat)

C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004

C;Accession: I84692; A46050

R;Pena de Ortiz, S.; Camon, M.M.; Jamieson, G.A.

Brain Res. Mol. Brain Res. 23, 278-283, 1994

A;Title: Expression of nuclear hormone receptors within the rat hippocampus: identification

A;Reference number: I48178; MUID:94335560; PMID:7914660

A;Accession: I84692

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-598 <PEN>

A;Cross-references: UNIPROT:Q07917; EMBL:U01146; NID:g515968; PIDN:AAC52143.1; PID:g5415

R;Seearce, L.M.; Laz, T.M.; Hazel, T.G.; Lau, L.F.; Taub, R.

J. Biol. Chem. 268, 8855-8861, 1993

A;Title: RNR-1, a nuclear receptor in the NGFI-B/Nur77 family that is rapidly induced in

A;Reference number: A46050; MUID:93232045; PMID:8473329

A;Accession: A46050

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-181, 'A', 183-249, 'PSP', 253-361, 'G', 363, 'DOCPROT', 372-433, 'S', 435-598 <SCE>

A;Cross-references: GB:L08595; NID:g310215; PIDN:AAA42058.1; PID:g310216

A;Experimental source: liver and brain

A;Note: sequence extracted from NCBI backbone (NCBIN:129782, NCBI:P.129783)

C;Genetics:

A;Gene: HZF-3

C;Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog

C;Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation; z

F;261-514/Domain: erba transforming protein homology <ERBA>

Query Match 51.8%; Score 1728; DB 2; Length 598;

Best Local Similarity 54.1%; Pred. No. 8.6e-88;

Matches 360; Conservative 74; Mismatches 124; Indels 108; Gaps 18;

Qy 1 MPCVQAQYSPSPGSSVAAQTY---SSEVTTIMAPDYTKLTWDLGSGTBITATITSLP 56

Db 1 MPCVQAQYSGSPQASPASQSYSHSGEYSSFLTFEFVKFMSDLTNTLIT--TATITSLP 58

Qy 57 S1STFVSGYSNTVELKPCSVYQM---QREPLIKVEGRAPSYHHHHHHHHHHHHHHHHH 112

Db 59 SFSTFMDNYSTGVDPKPCLYQWPLSQQSSIKVEDIQMHYQOHS----- 105

Qy 113 QQPSIPASSPDEVLTP-STSMYFKQSPPTPTTAPFPQAGALWDEALSPACCIAPGP 171

Db 106 ----LPPQS---EEMPHSGSVYKPPSPPTSTPSTFQVQHSPWDD-----PGS 148

Qy 172 LLDPPMKAVPT-----VAGARFPLPHFKSPHPPEA-----PSPAG 207

Db 149 LHNPHQNVATTHMIEQKTPVSRSLSPFSKQSPGTPGVSSCOMRFDGFLHVPNMPPEAG 208

Qy 208 GHHLGYDPTAAALSLPLGAAAAAGSQAALSHPHYGLP-LAKRAAPLAFPLGLTSPPT 266

Db 209 SHHV-----VDGQTFVAVNPPIRKPASMGFFGLQI---GH 239

Qy 267 ASSLLGESPSLPPSRSSSGEGTCVAGCDNAACQHYGVRTCEGCKGFFKRTVQKNKY 326

Db 240 ASQLL--DTQVPSPPSRGSPSNEGLCAVCGDNAACQHYGVRTCEGCKGFFKRTVQKNKY 297

Qy 327 VCLANKNCVDKRRNRRCQYCFQKCLSVGMVKEVVRTDSLKGRGRPLSPKPSLQOEP 386

Db 298 VCLANKNCVDKRRNRRCQYCFQKCLAVGMVKEVVRTDSLKGRGRPLSPKPS-----P 352

Qy 387 SOPSPPSPPICMNALVRALTDSTP--RDLDSRY--CPTDQAAAGTDAHVQOQFYNLLTA 443

Db 353 QDPSPPSPVSLISALVRAHVDSNPAMTSLDSYRFOANPDYQMSGDDTQHIOQFYDLLTG 412

Qy 444 SIDVRSWAEEKIPGFTDLPKEDQTLIESAFLEFVLRLSIRSNATDAEDKFFVFCNGLVLR 503

Db 413 SMEIIRGWAEEKIPGFADLPKADQDLIFESAFLEFVLRLAYRSNPVEGKLIFFCNGVLR 472

Qy 504 LQCLRGFGEWLSIDKDFSLNLSQSLNLDIOALACLSALSMTIRHGLKPKRVEELCNKIT 563

Db 473 LQCVRGFGEWLSIDKDFSLNLSQSLNLDIOALACLSALSMTIRHGLKPKRVEELCNKIT 532

Qy 564 SSKLDHQ--SKGOALEPTE--SKVIGALVELRKTCTGLQRIFFYLKLEDLVSPPSIIDKLF 620

Db 533 NCLKDHVTFNNGLNRPNYLSKLLGKLPRLTCTCQGLQRIFFYLKLEDLVPPPAIIDKLF 592

Qy 621 LDTLPF 626

Db 593 LDTLPF 598

RESULT 7

A37251

Probable nuclear hormone receptor NAK1 - human

N;Alternate names: early response protein NAK1; nuclear hormone receptor ST-59; steroid

C;Species: Homo sapiens (man)

C;Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 09-Jul-2004

C;Accession: A37251; A61191; A61451

R;Nakai, A.; Kartha, S.; Sakurai, A.; Toback, F.G.; DeGroot, L.J.

Mol. Endocrinol. 4, 1438-1443, 1990

A;Title: A human early response gene homologous to murine nur77 and rat NGFI-B, and rela

A;Reference number: A37251; MUID:91133413; PMID:2283997

A;Accession: A37251

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-598 <NAK>

A;Cross-references: UNIPROT:P22736; GB:D49728; NID:g1813881; PIDN:BAA08565.1; PID:g18138

R;Bondy, G.P.

Cell Growth Differ. 2, 203-208, 1991

A;Title: Phorbol ester, forskolin, and serum induction of a human colon nuclear hormone

A;Reference number: A61191; MUID:91329278; PMID:1651101

RESULT 8
A:Accession: A61191
A:Molecule type: mRNA
A:Residues: 35-252, 'L', 254-598 <BON>
R:Chang, C.; Kokontis, J.; Liao, S.; Chang, Y.
J. Steroid Biochem. 34, 391-395, 1989
A:Title: Isolation and characterization of human TR3 receptor: a member of steroid recep
A:Reference number: A61451; MUID:90173205; PMID:2626032
A:Accession: A61451
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-261, 'P', 263-369, 'L', 371-598 <CHA>
A:Cross-references: GB:L13740; NID:G292833; PIDN:AAA36763.1; PID:G292834
C:Genetics:
A:Gene: GDB:HMR
A:Cross-references: GDB:125357; OMIM:139139
A:Map position: 12q13-12q13
C:Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog
C:Keywords: DNA binding; hormone receptor; nucleus; transcription regulation; zinc finger
F:265-514/Domain: erba transforming protein homolog <ERBA>

Query Match 41.0%; Score 1367; DB 2; Length 598;
Best Local Similarity 47.5%; Pred. No. 5.7e-68;
Matches 317; Conservative 74; Mismatches 165; Indels 112; Gaps 23;

QY 1 MPCVQAQY---SPSPGSSVAAQYSSYETTEIMNPDYTKLTMDLGSSTEITATATSLPS 57
DB 1 MPCVQAQY---SPSPGSSVAAQYSSYETTEIMNPDYTKLTMDLGSSTEITATATSLPS 57
QY 58 ISTFVGYSNYELKPCVYQMO---RPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHH 114
DB 53 FSTFMDGYTGEFD---TFYQLPGTVPQCSSASSSTSSSATSPASASFKFDFQVY 109
QY 115 PSTP-PASPEDDEVLTSTMYFKQSPS--TPPTAP-PPQAG----- 153
DB 110 GCYFGLSPFVDEALSSGSDYYGSPCSAPSPSTPSPQPSGFWGSGFHSFSPSYEG 169
QY 154 -ALWDELALSPACCIAPGLLDPMKAVPTVAGARFPLFHFKPSPPHPPAPSPAGGHILG 212
DB 170 LRAWTQLPKASG-----PPQ-----PPAFSPSPPTGSPS----- 201
QY 213 YDTAAALSL-PLGAAAGSQAALSHYPYGLPLAKRAAPLAPPLGLTSPSTASLL 271
DB 202 ---LAQSPKLKPSQATHQLG-----EGESYSM-----PTAFPLG-----APTSPHLE 241
QY 272 G-----RSPSLPSPSSSSG--EGTCAVCGDNAAQHYGVRTCEGCKGFFKRTVQKNAK 325
DB 242 GSGILDTP-VTSTKARSGAPGSGRCVACGDNASQHYGVRTCEGCKGFFKRTVQKNAK 300
QY 326 YVCLANKNCVDRRRNRQYCEFOKCLSVGMVKEVVRTDLSKGRGRPLSPKPSKPLQOE 385
DB 301 YICLANKDCPDKRRNRQCFKQKCLAVGMVKEVVRTDLSKGRGRPLSPKPSKPLQOE 356
QY 386 PSQSPSPSPPICMNALVRALTDSTP--RDLSRYCPTDQAAAG--TDAEHVQOQYNILT 442
DB 357 -----PDASPNLLTSLVRAHLDSGSTAKLDYSKEQLVLPFHFGKEDAGDVQOQFDLLS 411
QY 443 ASIDVSRWAETKIPGTDLPKEDQTLIESAFLELFLVRLSIRSNTPAEKFFVFCNGLVLIH 502
DB 412 GSLEVRKWAETKIPGFAELSPADQDLLESAPLELFLVRLSIRSNTPAEKFFVFCNGLVLIH 471
QY 503 RLQCLRGFGLWDSIKDFSLNLSINLDIQALCLALSMTIRHGLKEPKRVEELCNKI 562
DB 472 RLQCARGFGLWDSILAFSLSLVLDVPAFACTLSALVITDRHGLQEPKRRVEELQNR 531
QY 563 TSSLKDHQKGOALEPTE---SKVLGALVELRKICTLGLQRTFYKLKLEDVSPSPIDK 618
DB 532 ASCLKEHVA-AVAGEPQASCLSLGLKLPKLCTCTQGLQRTFYKLKLEDVSPSPIDK 590
QY 619 LFLDTLPF 626
DB 591 IFMDTLFP 598

hormone receptor NGFI-B homolog - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
A:Accession: S33763; S31713
R:Smith, T.S.; Matharu, P.J.; Sweeney, G.E.
Biochim. Biophys. Acta 1173, 239-242, 1993
A:Title: Cloning and sequencing of a Xenopus homologue of the inducible orphan receptor
A:Reference number: S33763; MUID:9327961; PMID:8504173
A:Accession: S33763
A:Molecule type: mRNA
A:Residues: 1-577 <SMI>
A:Cross-references: UNIPROT:Q04913; EMBL:X70700; NID:G64916; PIDN:CAA50031.1; PID:G6491
C:Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog
C:Keywords: DNA binding; nucleus; transcription factor; zinc finger
F:244-493/Domain: erba transforming protein homolog <ERBA>

Query Match 40.5%; Score 1350.5; DB 2; Length 577;
Best Local Similarity 47.1%; Pred. No. 4.4e-67;
Matches 305; Conservative 88; Mismatches 162; Indels 93; Gaps 19;

QY 1 MPCVQAQYSPSPGSSVAAQYSSYETTEIMNPDYTKLTMDLGSSTEITATATSLPS 60
DB 1 MPCVQAQY---GSLSQACAGPCDNVVDILNSEFCGFTMDLVNSEL--AASTLSLPSFT 53
QY 61 FVEGYSSNYELKPCVYQMO---QRPILKVEGRAPSYHHHHHHHHHHHHHHHHHHH 116
DB 54 FMDGYTGEFD---AFYQLPSSNQSSSLKVE-----FQVFG 87
QY 117 IPAS--SPEDEVLTSTMYFKQSPS--TPPTAPPPQAGALWDELALSPACIARGPL 172
DB 88 CYPGFTNOLDTETMSGSDYYGSPCSIPSPSTPGQNPOLPTWE---CSYGAYSTQN 143
QY 173 LDP-----PMKAVPTVAGARFPLFHFKPSPPHPPAPSPAGGHILGYDPTAAALSLPLGAA 228
DB 144 YDMRHTWQKNSISQOTF--FSP-GTFAHSP-----NWAANPLKIATPHTLDQ 192
QY 229 AAGSQAAALSHYPYGLPLAKRAAPLAPPLGLTSPSTASSLLGESPSLPSPSSSSG 288
DB 193 LVDTDVFAALQNSSAGFPVAV-----PLGQAPGLDSSVLLDSPSPS-KTRSPSSN 242
QY 289 EGTCAVCGDNAAQHYGVRTCEGCKGFFKRTVQKNAKYCLANKNCVDRRRNRQYCR 348
DB 243 EGRCAVCGDNASQHYGVRTCEGCKGFFKRTVQKNAKYCLANKNCVDRRRNRQYCR 302
QY 349 FOKCLSVGMVKEVVRTDLSKGRGRPLSPKPSKPLQOEPSPSPPICMNALVRALTD 408
DB 303 FOKCLVGMVKEVVRTDLSKGRGRPLSPKPSKPLQOEPSPSPPICMNALVRALTD 353
QY 409 STPRD--LDYSRYCPTDQAAAGTDAE-HVQOQYNLLTASIDVSRWAETKIPGFTDLPKED 465
DB 354 SIPSSSKLDYSKFQETVPLQLEKSSVDVQOQFYDLSGSLEVRKWAETKIPGFTDLPKED 413
QY 466 QTLIESAFLELFLVRLSIRSNTPAEKFFVFCNGLVRLQCLRGFGLWDSIKDFSLNLO 525
DB 414 QDLLESFLELFLVRLSIRSNTPAEKFFVFCNGLVRLQCLRGFGLWDSIKDFSLNLO 473
QY 526 SLNLDIQALCLALSMTIRHGLKEPKRVEELCNKITSLLKDH-----OSKGOALEP 578
DB 474 RNWIDVSPSCLSALVITDRHGLKEPKRVEELQSQIINCLKEHPTSSWNEQRPNCLE-- 531
QY 579 TESKVLGALVELRKICTLGLQRTFYKLKLEDVSPSPIDKFLDTLPF 626
DB 532 --SKLGLKLPKLCTCTQGLQRTFYKLKLEDVSPSPIDKFLDTLPF 577

RESULT 9
ORMSN1
probable hormone receptor N10, nuclear - mouse
N:Alternate names: steroid hormone receptor homolog nur/77
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

C:Accession: S06953; A30059; A31344
R:Ryseck, R.P.; Macdonald-Bravo, H.; Mattei, M.G.; Ruppert, S.; Bravo, R.
EMBO J. 8, 3327-3335, 1989
A:Title: Structure, mapping and expression of a growth factor inducible gene encoding a
A:Reference number: S06953; MUID:90059925; PMID:2555161
A:Accession: S06953
A:Molecule type: DNA
A:Residues: 1-601 <RYS>
A:Cross-references: UNIPROT:P12813; EMBL:X16995; NID:G53312; PIDN:CAA34862.1; PID:G53313
R:Hazel, T.G.; Nathans, D.; Lau, L.F.
Proc. Natl. Acad. Sci. U.S.A. 85, 8444-8448, 1988
A:Title: A gene inducible by serum growth factors encodes a member of the steroid and th
A:Reference number: A94211; MUID:89042200; PMID:3186734
A:Accession: A30059
A:Molecule type: mRNA
A:Residues: 1-601 <HAZ>
A:Cross-references: GB:J04113; NID:G200115; PIDN:AAA39843.1; PID:G200116
C:Genetics:
A:Map position: 15P
A:Introns: 295/3; 339/1; 389/3; 457/2; 517/1
C:Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog
C:Keywords: DNA binding; nucleus; steroid hormone receptor; transcription factor; zinc f
F:268-517/Domain: erba transforming protein homology <ERBA>
F:270-290/Region: zinc finger CCCC motif
F:276-335/Domain: DNA binding #status predicted <DNA>
F:306-330/Region: zinc finger CCCC motif

Query Match 39.5%; Score 1316.5; DB 1; Length 601;
Best Local Similarity 44.9%; Pred. No. 3.4e-65;
Matches 307; Conservative 78; Mismatches 158; Indels 141; Gaps 22;

QY 1 MPCVQAQY-----SPSPGSSYAAQYSSSEYTEIMNPDYTKLTMDLGSSTEITATATLSL 55
DB 1 MPCIQAYGTATSPGP-----RDHLTGDPALAEFGKPTWDLASPETAAPATL 50

QY 56 PSISTFVEGYSNYELKPCVYQM---QRPLIKVEGRAPSTHHHHHHHHHHHHHQQOH 112
DB 51 PSFSTFMDGYTGFEFD---TFLYQLPGTGTQPCSSACSSASST-----SSS 91

QY 113 QQPSTPPASS-----PEDEVLPSTSMYFKQSPPS--TPTTAPFPQ 151
DB 92 SSSATSPASASPKFEDFQVYCYPGTISGLPDETSSSGSEYGFPCSPSTNFQPS 151

QY 152 AGALWD-----EALPSAPGCIAPGLLDPMPKAVPTVAGARFPLFH 192
DB 152 QLSPWDGSGFHFSPSTYEGLWATEQLPKA---SSGP---PP-----PPTF 192

QY 193 FKPSPPHPAPSPAGGHHLGVDPTAAALSLPLGAAAAGSQAALESHYPGLPLAKRAA 252
DB 193 FFSFPTGSPSLAQSLLKLPFPATHQLG-----EGESYSM----- 229

QY 253 PLAFPLGLT-PSPTASSLGSPLSPSPSSSSG--EGTCVACGDNAACQHYGVRTC 309
DB 230 PAAPFGLATSPNRTSGIL-DAP-VTSKSRSGASGGSEGCACVCGDNASCQHYGVRTC 287

QY 310 EGCKGFFKRTVQNAKYVCLANKNCPVDKRRNRNQYCRFKCLSLGVMYKEVVRTDSLKG 369
DB 288 EGCKGFFKRTVQNAKYVCLANKNCPVDKRRNRNQYCRFKCLAVGMVKEVVRTDSLKG 347

QY 370 RGRPLSPKPSKPLQOEPSQSPSPPICMNALVRALTDSTP--RDLDYSRYCPTDQAAA 427
DB 348 RGRPLSPKPKQP-----PDASPTNLLTSLIRAHLDSPSTAKLDYSKQELVLPF 398

QY 428 G-TDAHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLESFAFELFVLRLSRS 486
DB 399 GKHEDAGDVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLESFAFELFVLRLSRS 458

QY 487 NTAEDKVFVFCNGLVLRHLOCLRGFGWLDSDIKDFSLNQLSLNDIQAALCSALSMTTER 546
DB 459 KPGEGLKIFCSGLVLHOLQCARFGDWIDNILAFSRSLHSLGVDVPAFACLSALVLITDR 518

QY 547 HGLKEPKRVVELCNKLTSSLKDHOSKGOALEPTE-----SKVLGALVELRKICTLGLQRI 602
DB 547 HGLKEPKRVVELCNKLTSSLKDHOSKGOALEPTE-----SKVLGALVELRKICTLGLQRI 602

DB 519 HGLQDPRRVELQNRNASCLAEHWAT-VAGDPQAPQASCLSRLLGKLPELRLCTQLQRI 577

QY 603 YLKLEDLVSPSPSIIDKFLDLTLPF 626
DB 578 CLKLEDLVPPPPPIVDKIFMDTLSP 601

RESULT 10
JQ0623
Nerve growth factor-inducible protein B - rat
N:Alternate names: glucocorticoid receptor homolog
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Aug-1999
A:Accession: JQ0623; A34097
R:Milbrandt, J.
Neuron 1, 183-188, 1988
A:Title: Nerve growth factor induces a gene homologous to the glucocorticoid receptor g
A:Reference number: JQ0623; MUID:90166506; PMID:3272167
A:Accession: JQ0623
A:Molecule type: mRNA
A:Residues: 1-563 <MIL>
A:Cross-references: GB:U17254; NID:G596053; PIDN:AAA56770.1; PID:G596054
A:Experimental source: cell line PCL2
R:Watson, M.A.; Milbrandt, J.
Mol. Cell. Biol. 9, 4213-4219, 1989
A:Title: The NGFI-B gene, a transcriptionally inducible member of the steroid receptor g
A:Reference number: A34097; MUID:90066426; PMID:2479823
A:Accession: A34097
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 255-260; 298-303; 349-354; 417-422; 476-481 <WAT>
C:Comment: This protein is proposed to be a ligand-dependent transcriptional activator w
C:Genetics:
A:Gene: NGFI-B
C:Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homology
C:Keywords: DNA binding; growth factor receptor; nucleus; transcription regulation; zinc
F:230-479/Domain: erba transforming protein homology <ERBA>

Query Match 38.3%; Score 1277; DB 2; Length 563;
Best Local Similarity 46.6%; Pred. No. 4.7e-63;
Matches 297; Conservative 70; Mismatches 145; Indels 126; Gaps 20;

QY 40 MDLSTETATATSLPSISTFVE-GYSSNVELKPCVYQMQRPLIKVEGRAPSYHHHH 98
DB 1 MDLSPETAPTATPLSFSTFMDGGYTGFEFD---TFLYQLPGTGAQPCSSASST----- 52

QY 99 HHHHHHHHHQOQHQOQPSIPPASS-----PEDEVLPSTSMYFKQSP 139
DB 53 -----SSSSSATSPASASPKFEDFQVYCYPGTISGLPDETILSSGSDYIGSP 100

QY 140 PSTPT--TPAPPPQAGALWD-----EALPSAPGCIAPGLLDPMPK 178
DB 101 CSAPSPPTNPQPSQLSPWDGSGFHFSPSTYEGLRVWTEQLPKASG-----PPP-- 150

QY 179 AVPTVAGARFPLFHPKSPHPHPADSPAGGHHLGVDPTAAALSLPLGAAAAGSQAAL 238
DB 151 -----PPTTFSPSPPTGSPS-----LAQSSKL---FPAPATHQLGSG 186

QY 239 ESHPYGLPLAKRAAPLAFPPGLT-PSPTASSLGSPLSPSPSSSSG--EGTCVAC 295
DB 187 ESY-----SVPAAPFGLAPTSNCDTSGIL-DAP-VTSKARSSGSSGSEGCACV 235

QY 296 GDNAACQHYGVRTCBGCKGFFKRTVQNAKYVCLANKNCPVDKRRNRNQYCRFKCLSLV 355
DB 236 GDNASQHYGVRTCBGCKGFFKRTVQNAKYVCLANKNCPVDKRRNRNQYCRFKCLAV 295

QY 356 GMVKEVVRTDSLKGRGRGLPSKPSKPLQOEPSQSPSPPICMNALVRALTDSTPR--D 413
DB 296 GMVKEVVRTDSLKGRGRGLPSKPSKPLQOEPSQSPSPPICMNALVRALTDSTPR--D 413

QY 414 LDYSRYCPTDQAAAAG-TDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLES 472
DB 347 LDYSRYCPTDQAAAAG-TDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLES 406

A:Title: The nuclear receptor superfamily has undergone extensive proliferation and divergence
A:Reference number: Z22443; MUID:99148134; PMID:10022975
A:Accession: T43348
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-619 <SLU>
A:Cross-references: UNIPROT:P41829; EMBL:AF083224; NID:94139073; PIDN:AA03682.1; PID:94139073
R:Lightning, J.
submitted to the EMBL Data Library, August 1994
A:Reference number: Z19215
A:Accession: T20039
A:Molecule type: DNA
A:Residues: 58-619 <WIL>
A:Cross-references: EMBL:Z36237; PIDN:CAA85271.1; GSPDB:GN00021; CESP:C48D5.1
A:Experimental source: clone C48D5
R:Kostrouch, Z.; Kostrouchova, M.; Rall, J. E.
Proc. Natl. Acad. Sci. U.S.A. 92, 156-159, 1995
A:Title: Steroid/thyroid hormone receptor genes in *Caenorhabditis elegans*.
A:Reference number: A55595; MUID:95116514; PMID:7816808
A:Accession: I45068
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 157-187, 'r', 189-619 <RES>
A:Cross-references: EMBL:U13076; NID:9538372; PIDN:AAA96984.1; PID:9538373
C:Genetics:
A:Gene: nhr-6; cnr8; CESP:C48D5.1
A:Map position: 3
A:Introns: 71/1; 98/3; 124/3; 175/2; 240/2; 337/1; 389/3; 510/1; 550/3
C:Superfamily: *Caenorhabditis elegans* nuclear steroid hormone receptor NHR-6; *erba* transmembrane protein
C:Keywords: steroid hormone receptor; zinc finger
F:266-507/Domain: *erba* transmembrane protein homology <ERBA>

Query Match 16.6%; Score 554.5; DB 1; Length 619;
Best Local Similarity 28.4%; Pred. No. 2.4e-23;
Matches 190; Conservative 79; Mismatches 249; Indels 151; Gaps 24;
QY 4 VQAQVSPSPGSSYAAQTVSYVTEIMNPDTYKLTMDLGSFTEATATTSPLSTFVE 63
DB 11 LQDQFNCSPAS--VDSSYSSCSVEDEIEIYTLVRN-----EEPLRDRFFR 56
QY 64 GYSSNYELPSCVY-----QWQPLIKVEGRAPSYHHHHHHHHHHHHHHHQQQPSI 117
DB 57 ESKNSSCSSFDYGFSGSSSRKSKTDDADLSL-----FHLVTSQVNTV 107
QY 118 PRASSPEDEVLSTSMYFQSPSTPTTPAPFPQAGALWDEALPSAPGCIAP--GFLDLP 175
DB 108 PRPTKTEVESIPEE---FQKPSSS-----SHRLPSEMNASITHIKSELDP 150
QY 176 PMKA-----VPTVAGARPLF---HFKSPSPHPAPSPAGGHLGYDPTAAALSL 223
DB 151 TMOAFQMPHNDLFLATAAPHYNPFALSNDFMPLMPSFTSPFPQH-----F 198
QY 224 PLGAAAAGSQAALSHYPGLPLAKRAAPLAPPLGLTPSPASSLLGESP----- 275
DB 199 FV-SDRRSGSQGTTSSN-----NTGGTPSPHSSSLTPSPLOQFLRS 241
QY 276 -----SLPSP-----PSRSSSGEGTCVCGNNAACQHYGVRTCEGKGFRTVQKNAK 325
DB 242 FLNPDNLSTPTSGVSEFSETALDADKMCVNCNDRVAVCLHYGARTCEGCKGFRTVQKNSK 301
QY 326 YVCLANKNCPVQRRNRNCOYCRFOKCLSVGMVKEVVRTDLSKGRGRLPSPKSPLOQE 385
DB 302 YTCAGNKTCPIDKRRYRSRQYCRQKCLEVGMVKEIVRHGSLSGRRGLSSKTK--LARS 359
QY 386 PSQSPSPSPICMNNALVRALTSTPRDLDSRYCPTDQAAAGTDAEHVQQFVNLITASI 445
DB 360 EDQSPSPPLALMG---KAIEDHTNMTVVRQFMQPFDEITIA-----LRLHGL 406
QY 446 DVRSWAEEKIPGTDLPKEDQTLIESAFLEFLVRLSIRSNATDAEDKVFVCLNGLVHLRQ 505
DB 407 HATKLLMAMPQISEIQPADFQILLSRSPFAIMAIRVANRCNSTDTIMFESG-----ELF 462
QY 506 CLRGFGE-----WLDISKDFSLNQLSLNLDIOALACLALSMTI---TERH--GLKEP 552

DB 463 SLNAPPACFOQIRFMVDMKARTFS---SLVDWEPQAFAPFIALQFLAGNTHVNLGLTNK 519
QY 553 KRVELCNKLTSSLSKDHQSGQALEPTESKVGALVELRKICTGLQRIFFYLKLEDLVSP 612
DB 520 PLVDQVOSTINAKDKHCSGSQ---NKLAKIVRLTQDFVDFHALGLQAL-----DILYP 570
QY 613 PSIIDKLF 621
DB 571 SHQLPEEFM 579
RESULT 14
G88408
protein nhr-6 [imported] - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-May-2002
C:Accession: G88408
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biological processes
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999
A:Accession: G88408
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-562 <STO>
A:Cross-references: GB:chr_III; PIDN:CAA85271.1; PID:G3875028; GSPDB:GN00021
C:Genetics:
A:Gene: nhr-6
A:Map position: 3
C:Superfamily: *Caenorhabditis elegans* nuclear steroid hormone receptor NHR-6; *erba* transmembrane protein

Query Match 16.6%; Score 554; DB 2; Length 562;
Best Local Similarity 30.1%; Pred. No. 2.3e-23;
Matches 169; Conservative 72; Mismatches 199; Indels 122; Gaps 20;
QY 105 HHHHQHQHQPSIPASSPEDEVLSTSMYFQSPSTPTTPAPFPQAGALWDEALPSAP 164
DB 38 HSLVTSQVNTVPKPTKTEVESIPEE---FQKPSSS-----SHRLPSEM 80
QY 165 GCIAP--GFLDLP 210
DB 81 NASITHIKSELDTMOAFQMPHNDLFLATAAPHYNPFALSNDFMPLMPSFTSPFPQH 140
QY 211 LGYDPTAAALSLPLGAAAAGSQAALSHYPGLPLAKRAAPLAPPLGLTPSPASSLL 270
DB 141 -----FV-SDRRSGSQGTTSSN-----NTGGTPSPHSSSL 171
QY 271 LGESP-----SLPSP-----PSRSSSGEGTCVCGNNAACQHYGVRTCEGC 312
DB 172 PTSPPLOQFLRSFLNPDNLSTPTSGVSEFSETALDADKMCVNCNDRVAVCLHYGARTCEGC 231
QY 313 KGFRTVQKNAKYVCLANKNCPVQRRNRNCOYCRFOKCLSVGMVKEVVRTDLSKGRRG 372
DB 232 KGFRTVQKNSKYTCAGNKTCPIDKRRYRSRQYCRQKCLEVGMVKEIVRHGSLSGRRG 291
QY 373 RLPSKPSPLQOEPSQSPSPPICMNNALVRALTSTPRDLDSRYCPTDQAAAGTDAE 432
DB 292 RLSSKTK--LARSEDPSPPLALMG---KAIEDHTNMTVVRQFMQPFDEITIA----- 341
QY 433 HVQPFYNLLTASIDVRSWAEEKIPGTDLPKEDQTLIESAFLEFLVRLSIRSNATDAEDK 492
DB 342 -----LRLHGLHATKLLMAMPQISEIQPADFQILLSRSPFAIMAIRVANRCNSTDT 396
QY 493 FVFCNGLVHLRQCLRGFGE-----WLDISKDFSLNQLSLNLDIOALACLALSMTI- 543
DB 397 IMFESG-----ELFSLNAPFACFOQIRFMVDMKARTFS---SLVDWEPQAFAPFIALQ 449
QY 544 ---TERH--GLKEPKEVEELCNKLTSSLSKDHQSGQALEPTESKVGALVELRKICTGLQ 599
DB 450 GNTHEVNLGLTNKPLVDQVOSTINAKDKHCSGSQ---NKLAKIVRLTQDFVDFHALGLQ 506

Job time : 44 secs

QY 600 RIFYLKLELVSPPSIIDKLEL 621
Db 507 AL-----DILYPSHQLPEEFM 522

RESULT 15
SI3512
retinoic acid receptor beta-1 - chicken
C:Species: Gallus gallus (Chicken)
C:Date: 21-Nov-1993 #sequence_revision 22-Apr-1995 #text_change 09-Jul-2004
C:Accession: SI3512; SI7073
R:Padanilam, B.J.; Mcleod, L.B.; Suzuki, H.; Solursh, M.
Nucleic Acids Res. 19, 395, 1991
A:Title: Nucleotide sequence of an isoform of chicken retinoic acid binding protein-beta
A:Reference number: SI3512; MUID:91195065; PMID:1849630
A:Accession: SI3512
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-455 <PAD>
A:Cross-references: UNIPROT:P22448; EMBL:X56674; NID:G63785; PIDN:CAA39997.1; PID:G63786
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1990
R:Nohno, T.; Muto, K.; Noji, S.; Saito, T.; Taniguchi, S.
Biochim. Biophys. Acta 1089, 273-275, 1991
A:Title: Isoforms of retinoic acid receptor beta expressed in the chicken embryo.
A:Reference number: SI6243; MUID:91274365; PMID:1647216
A:Accession: SI7073
A:Molecule type: mRNA
A:Residues: 1-29, 'P', 31-80 <NOH>
C:Superfamily: retinoic acid receptor alpha; erba transforming protein homology
C:Keywords: alternative splicing; DNA binding; nucleus; transcription regulation; zinc f
F:86-337/Domain: erba transforming protein homology <ERBA>
F:88-108/Region: zinc finger
F:124-148/Region: zinc finger

Query Match 16.3%; Score 543; DB 2; Length 455;
Best Local Similarity 32.5%; Pred. No. 7 4e-23;
Matches 149; Conservative 65; Mismatches 171; Indels 74; Gaps 14;

QY 180 VPTVAGARPLPHFKPSPHPAPSPAGGHILGYDPTAAALSLPLGAAAAAGSQAAALE 239
Db 10 VPAVNG---HMTHY-PAAPYPLLFAPIG-----GLSLP-----SLHGLQ 45

QY 240 SHPYGLPLAKRAAPLAFPLGL-TPSP-----TASSLLGESPSLSPSPSSSSSGEGT 291
Db 46 SHP-----PTSGCSTPSATVETQSTSESELVSPSPSPPLPPRVYK----P 87

QY 292 CAVCGDNAACOHVGRTCGCGFFKRTYQKNKAKVYCLANKNCVDRKRNRNCOYCRFOK 351
Db 88 CFVCQDKSSGYHYGVSAACGCGKGFRRSIQKNVYTCRDKKNCVINKVTRNRCQYCRLOK 147

QY 352 CLSVGMVKEVVRVTDLSLKGRRRLPSKPSGLOQEPSQSPSPSPPICMNNALVRALTDSTP 411
Db 148 CFVEVGMKESVNRNK-----KKKEPTKQESTENYEMTAELDDLTEKIRKAHQET- 198

QY 412 RDLVSVRYCPTQAAAGTDAEH-----VOQFYNLLTASIDVSRSWAEKIPGFTDLPK 463
Db 199 ----FPSLCQLGKYTTNSADHRVRLDLGLWDKFSLELATKCIKIVEFAKRLPGFTSLTI 254

QY 464 EDQTLIESAFLEFLVRLSIRNTAEDKFVFCNGLVLAHRLQCLR-GFGEWLDSIKDFSL 522
Db 255 ADQITLLKAACLDILRLICTRYTPEQDTMTFSQDTNLTNRQMNAGFGPLDVLVFTFAN 314

QY 523 NQQLNLMDITQALACLSALSMIT-ERHGLKEPKRVEELCNKITSSLNKHQSGQALEP-TE 580
Db 315 QLLPLEMDDETGLLSAICLICDRQDRLDEPKVKYDQLEPFLLEALKIYIRKRPKNKPHMF 374

QY 581 SKVLGALVELRKICTLGLQRIFYFLKLEDLVSPPSIIDKL 619
Db 375 PKILMKITDLRSISAKGAERVITLKWEIFGSMPLIQEM 413

Search completed: March 12, 2005, 17:52:00

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2005, 17:40:16 ; Search time 176 Seconds
(without alignments)
1821.373 Million cell updates/sec

Title: US-10-608-863-2
Perfect score: 3337
Sequence: 1 MPCVQAQSPSPGSSYAAQ.....EDLVSPPSIIDKFLDLPF 626

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	ID	Description
1	3333	99.9	626	1 NR43 HUMAN	Q92570 homo sapien
2	3133.5	93.9	643	2 Q97726	Q97726 sus scrofa
3	3074	92.1	628	1 NR43 RAT	P51179 rattus norv
4	3054.5	91.5	627	1 NR43 MOUSE	Q9qzb6 mus musculu
5	2099.5	62.9	446	2 Q97727	Q97727 sus scrofa
6	1751	52.5	598	1 NR42 HUMAN	P43354 homo sapien
7	1746	52.3	598	1 NR42 MOUSE	Q06219 mus musculu
8	1736	52.0	598	1 NR42 RAT	Q07917 rattus norv
9	1677	50.3	600	2 Q98TQ3	Q98TQ3 oryzias lat
10	1588	47.6	586	2 Q6DH08	Q6dh08 brachydanio
11	1572	47.1	535	2 Q6NXU0	Q6nxu0 homo sapien
12	1371	41.1	598	1 NR41 CANFA	P51666 canis famlil
13	1368	41.0	598	2 Q8N3V2	Q8n3v2 homo sapien
14	1367	41.0	598	1 NR41 HUMAN	P22736 homo sapien
15	1364	40.9	598	2 Q61BU8	Q61bu8 homo sapien
16	1353	40.5	652	2 Q6ZMM6	Q6zmm6 homo sapien
17	1351.5	40.5	577	2 Q61NY4	Q6iny4 xenopus lae
18	1350.5	40.5	577	1 NR42 XENLA	Q04913 xenopus lae
19	1350.5	40.5	577	2 Q7T0V3	Q7t0v3 xenopus lae
20	1322.5	39.6	597	1 NR41 RAT	P22829 rattus norv
21	1316.5	39.5	601	1 NR41 MOUSE	P12813 mus musculu
22	1254	37.6	574	2 Q6GMG3	Q6gmG3 brachydanio
23	1219.5	36.5	808	2 Q704C6	Q704c6 anopheles g
24	1197.5	35.9	545	2 Q9U4L1	Q9u4l1 aedes aegypt
25	1190.5	35.7	1073	1 HR38 DROME	P49869 drosophila
26	1168.5	35.0	556	2 Q81NU7	Q81nu7 drosophila
27	1057.5	31.7	392	2 Q9R1W4	Q9r1w4 mus musculu
28	1023	30.7	360	1 HR38 BOMMO	P49870 bombyx mori
29	554.5	16.6	619	1 NR36 CAEEL	P41829 caenorhabdi
30	545	16.3	455	1 RRB_CHICK	P22448 gallus gall
31	545	16.3	455	1 RRB_COTJA	Q9w6b3 coturnix co

32	541	16.2	156	2 Q6AW80	Q6aw80 homo sapien
33	541	16.2	482	1 RRB_MOUSE	P22605 mus musculu
34	533.5	16.0	533	1 RXRB_HUMAN	P28702 homo sapien
35	531.5	15.9	462	1 RRA_MOUSE	P11416 mus musculu
36	531.5	15.9	462	2 P97513	P97513 mus spretus
37	531	15.9	455	1 RRB_HUMAN	P10826 homo sapien
38	530.5	15.9	462	1 RRA_HUMAN	P10276 homo sapien
39	529.5	15.9	525	2 Q95L53	Q95l53 mustela vis
40	527	15.8	590	2 Q6P2H5	Q6p2h5 homo sapien
41	527	15.8	600	2 Q6IN48	Q6in48 homo sapien
42	523.5	15.7	460	1 RRA_CHICK	Q90966 gallus gall
43	521.5	15.6	520	1 RXRB_MOUSE	P28704 mus musculu
44	520.5	15.6	458	1 RRA_NOTVI	P18514 notophthalm
45	519.5	15.6	451	2 Q6MGB3	Q6mgb3 rattus norv

ALIGNMENTS

RESULT 1
NR43_HUMAN
ID NR43_HUMAN STANDARD; PRT; 626 AA.
AC Q92570; Q12935; Q14979; Q16420; Q9UEK2; Q9UEK3;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Orphan nuclear receptor NR4A3 (Nuclear hormone receptor NOR-1)
DE (Neuron-derived orphan receptor 1) (Mitogen induced nuclear orphan receptor).
DE receptor.
GN Name=NR4A3; Synonyms=CHN, CSMF, MINOR, NOR1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC TISSUE=Fetal brain;
RX MEDLINE=96404972; PubMed=8809112; DOI=10.1016/0167-4781(96)00101-7;
RA Ohkura N., Ito M., Tsuchida T., Sasaki K., Yamaguchi K., Miki K.;
RT "Structure, mapping and expression of a human NOR-1 gene, the third member of the Nur77/NGFI-B family.";
RL Biochim. Biophys. Acta 1308:205-214(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC TISSUE=Periphereal blood;
RX MEDLINE=96192925; PubMed=8614405; DOI=10.1210/me.9.12.1692;
RA Hedvat C.V., Irving S.G.;
RT "The isolation and characterization of MINOR, a novel mitogen-inducible nuclear orphan receptor.";
RL Mol. Endocrinol. 9:1692-1700(1995).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC TISSUE=Fetal brain;
RX MEDLINE=96152889; PubMed=8570200;
RA Clark J., Benjamin H., Gill S., Sidhar S., Goodwin G., Crew J., Gusterson B.A., Shipley J., Cooper C.S.;
RT "Fusion of the EWS gene to CHN, a member of the steroid/thyroid receptor gene superfamily, in a human myxoid chondrosarcoma.";
RL Oncogene 12:229-235(1996).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM BETA), AND CHROMOSOMAL TRANSLOCATION WITH EWS.
RC TISSUE=Fetal heart;
RX MEDLINE=96177652; PubMed=8634690;
RA Labelle Y., Zucman J., Stenman G., Kindblom L.-G., Knight J., Turc-Carel C., Dockhorn-Dworniczak B., Mandahl N., Desmaziere C., Peter M., Aurias A., Delattre O., Thomas G.;
RT "Oncogenic conversion of a novel orphan nuclear receptor by chromosome translocation.";
RL Hum. Mol. Genet. 4:2219-2226(1995).
RN [5]
RP SEQUENCE OF 1-69 AND 301-443 FROM N.A. (ISOFORM BETA), AND ALTERNATIVE SPLICING.

FT	CONFLICT	240	240	G -> S (in Ref. 3 and 4).
FT	CONFLICT	454	454	K -> R (in Ref. 1).
FT	CONFLICT	579	579	T -> N (in Ref. 1).
FT	CONFLICT	585	585	G -> V (in Ref. 1).
SQ	SEQUENCE	626 AA;	68199 MW;	40D18DA8FEF9918C CRC64;

Query Match		99.9%;	Score 3333;	DB 1;	Length 626;
Best Local Similarity		99.8%;	Pred. No. 1.4e-168;		
Matches 625;	Conservative	0;	Mismatches	1;	Indels
Qy	1	MPCVQAQSPSPSSSYAAQYSSSEYTEIMNPDYTKLTMDLGSTETATATTSLSISIST 60			
Db	1	MPCVQAQSPSPSSSYAAQYSSSEYTEIMNPDYTKLTMDLGSTETATATTSLSISIST 60			
Qy	61	FVEGYSSNYELKPCSCVYQMQRPPLIKVEGRAPSVHHHHHHHHHHHHHHHHHHHHHHHHH 120			
Db	61	FVEGYSSNYELKPCSCVYQMQRPPLIKVEGRAPSVHHHHHHHHHHHHHHHHHHHHHHHHH 120			
Qy	121	SSPEDEVLPSTSMYFKQSPSPSTPTTPAFPQAGALMDEALPSAPGCIAPGFLDPPMKAV 180			
Db	121	SSPEDEVLPSTSMYFKQSPSPSTPTTPAFPQAGALMDEALPSAPGCIAPGFLDPPMKAV 180			
Qy	181	PTVAGARFPLFHFKEPSPHPAPSPAGCHHLYGVDPTAAALSLPLGAAAAGSQAALLEG 240			
Db	181	PTVAGARFPLFHFKEPSPHPAPSPAGCHHLYGVDPTAAALSLPLGAAAAGSQAALLEG 240			
Qy	241	HPYGLPLAKRAAPLAFPLGLTPTSPATSSLLGSPSPSPSSSSSGEGTCACVCGDNAA 300			
Db	241	HPYGLPLAKRAAPLAFPLGLTPTSPATSSLLGSPSPSPSSSSSGEGTCACVCGDNAA 300			
Qy	301	COHYGVRTCEGCKGFFKRTVQKNAYKVCNQCPCVDKRRNRRCQYCFQKCLSVGMVKE 360			
Db	301	COHYGVRTCEGCKGFFKRTVQKNAYKVCNQCPCVDKRRNRRCQYCFQKCLSVGMVKE 360			
Qy	361	VVRTDSLKGRGRRLPSPKPSQLOEQSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420			
Db	361	VVRTDSLKGRGRRLPSPKPSQLOEQSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420			
Qy	421	PTDQAAAGTDAEHVQOQFNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFLV 480			
Db	421	PTDQAAAGTDAEHVQOQFNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLEFLV 480			
Qy	481	RLSIRSNTAEDKVFVFCNGLVLRQLGRFGEWLDISKFSLNLQSLNLDIQALACLAL 540			
Db	481	RLSIRSNTAEDKVFVFCNGLVLRQLGRFGEWLDISKFSLNLQSLNLDIQALACLAL 540			
Qy	541	SMITRHGCKEPKRVHEELCNKITSSLDKDHOSKQGALEPTESKVLGALVELRKICTGLOR 600			
Db	541	SMITRHGCKEPKRVHEELCNKITSSLDKDHOSKQGALEPTESKVLGALVELRKICTGLOR 600			
Qy	601	IFYLKLEDLVSPPSIIDKFLDLPFP 626			
Db	601	IFYLKLEDLVSPPSIIDKFLDLPFP 626			

RESULT 2	
O97726	PRELIMINARY;
ID O97726	PRELIMINARY; PRT; 643 AA.
AC O97726;	
DT 01-MAY-1999 (TREMBLrel. 10, Created)	
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)	
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)	
DE Neuron-derived orphan receptor-1 alfa.	
GN Names-NOR-1;	
OS Sus scrofa (Pig).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.	
NCBI_TaxID=9823;	
RN [1]	
RP SEQUENCE FROM N.A.	
RA Cases-Langhoff C.; Castello A.; Martinez-Gonzalez J.; Badimon L.;	
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.	
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).	

```

CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL: AJ011767; CA009763.1; -.
DR HSP: P43354; IOVL.
DR GO: GO:0005634; C:nucleus, IEA.
DR GO: GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR000536; Hrmn_recept_lig.
DR InterPro: IPR003072; NucR_receptor.
DR InterPro: IPR003070; NucR_receptor.
DR InterPro: IPR008946; Str_ncl_receptor.
DR InterPro: IPR001723; Stdhrmn_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; Hormone_recep; 1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR01286; NORNURCEPTR.
DR PRINTS: PR01284; NUCLEARRECPTR.
DR PRINTS: PR00398; STRODHORMONER.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; Znf_C4steroid; 1.
DR SMART: SM00430; HOLI; 1.
DR SMART: SM00399; Znf_C4; 1.
DR PROSITE: PS00031; NUCLEAR RECEPTOR; 1.
DR DNA-binding: Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 643 AA; 70077 MW; 20596A4222FB4CA7 CRC64;

Query Match 93.9%; Score 3133.5; DB 2; Length 643;
Best Local Similarity 92.4%; Pred. NO. 5e-158;
Matches 594; Conservative 10; Mismatches 22; Indels 17; Gaps 3;

QY 1 MPCVQAYSPSPGSSYAAQTYSSEYTTIMPDYTKLTMDLGSITATATSLPSIST 60
DB 1 MPCVQAYSPSPGSSYAAQTYSSEYTTIMPDYTKLTMDLGSITATATSLPSIST 60

QY 61 FVEGYSSNVELKSPCVVQMO---RPLIKVEGRAPSYHHHHHHHHHHHHHQQ----- 110
DB 61 FVEGYSSNVELKSPCVVQMO---RPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHQQ----- 110

QY 111 --HQQPSIPASSPEDEVLPSMTMYFKQSPSTPTTAPPPQAGALWDEALPSAPGCTA 168
DB 111 --HQQPSIPASSPEDEVLPSMTMYFKQSPSTPTTAPPPQAGALWDEALPSAPGCTA 168

QY 121 PQQQQPSIPSPGPEDEVLPSMTMYFKQSPSTPTTTPVPQAGALWEDALPSAGGCTA 180
DB 121 PQQQQPSIPSPGPEDEVLPSMTMYFKQSPSTPTTTPVPQAGALWEDALPSAGGCTA 180

QY 169 PGFLDPPMKAVPTVAGAPFLFHFKPSPPHPAPSPAGGHHLYGDPYTAALSLPLG-- 226
DB 169 PGFLDPPMKAVPTVAGAPFLFHFKPSPPHPAPSPAGGHHLYGDPYTAALSLPLG-- 226

QY 181 PGFLDPPMKAVPTVAGAPFLFHFKPSPPHPAPSPAGGHHLYGDPYTAALSLPLGAA 240
DB 181 PGFLDPPMKAVPTVAGAPFLFHFKPSPPHPAPSPAGGHHLYGDPYTAALSLPLGAA 240

QY 227 ---AAAAGSQAALESHPVGLPLAKRAAPLAPPLGLTPSPASSLIGESPSPSPSR 283
DB 227 ---AAAAGSQAALESHPVGLPLAKRAAPLAPPLGLTPSPASSLIGESPSPSPSR 283

QY 241 AAAAAAGSQAALESHPVGLPLAKRAAPLAPPLGLTPSPASSLIGESPSPSPSR 300
DB 241 AAAAAAGSQAALESHPVGLPLAKRAAPLAPPLGLTPSPASSLIGESPSPSPSR 300

QY 284 SSSSGSGTCAVCGDNNAACQHYGVRTCEGCKGFFKRTVQKNKAVCLANKNCVDRKRRNR 343
DB 284 SSSSGSGTCAVCGDNNAACQHYGVRTCEGCKGFFKRTVQKNKAVCLANKNCVDRKRRNR 343

QY 301 STASGSGTCAVCGDNNAACQHYGVRTCEGCKGFFKRTVQKNKAVCLANKNCVDRKRRNR 360
DB 301 STASGSGTCAVCGDNNAACQHYGVRTCEGCKGFFKRTVQKNKAVCLANKNCVDRKRRNR 360

QY 344 CQYCRFQKCLSGVMKVEVTRTDSLKGRRGLPSKPSPLQQSPSPSPSPSPSPSPSPSP 403
DB 344 CQYCRFQKCLSGVMKVEVTRTDSLKGRRGLPSKPSPLQQSPSPSPSPSPSPSPSPSP 403

QY 361 CQYCRFQKCLSGVMKVEVTRTDSLKGRRGLPSKPSPLQQSPSPSPSPSPSPSPSPSP 420
DB 361 CQYCRFQKCLSGVMKVEVTRTDSLKGRRGLPSKPSPLQQSPSPSPSPSPSPSPSPSP 420

QY 404 RALTDSTPRDLVSRYCPTDQAAGTDAEHVQOYFNLLTASIDVSSSWAEKIPGFTDLPK 463
DB 404 RALTDSTPRDLVSRYCPTDQAAGTDAEHVQOYFNLLTASIDVSSSWAEKIPGFTDLPK 463

QY 421 RALTDSTPRDLVSRYCPTDQAAGTDAEHVQOYFNLLTASIDVSSSWAEKIPGFTDLPK 480
DB 421 RALTDSTPRDLVSRYCPTDQAAGTDAEHVQOYFNLLTASIDVSSSWAEKIPGFTDLPK 480

QY 464 EQDTLLIESAFLEFVLRISIRNTAEDKVFVFCNGLVLRHLOCLRGFGWLDSIKDPSLN 523
DB 464 EQDTLLIESAFLEFVLRISIRNTAEDKVFVFCNGLVLRHLOCLRGFGWLDSIKDPSLN 523

QY 481 EQDTLLIESAFLEFVLRISIRNTAEDKVFVFCNGLVLRHLOCLRGFGWLDSIKDPSLN 540
DB 481 EQDTLLIESAFLEFVLRISIRNTAEDKVFVFCNGLVLRHLOCLRGFGWLDSIKDPSLN 540

QY 524 LOSNLNDIQALACLSALSMITERHGLKEPKRVEELCNKITSSLKDHQSGQALEPTESKV 583
DB 524 LOSNLNDIQALACLSALSMITERHGLKEPKRVEELCNKITSSLKDHQSGQALEPTESKV 583

QY 541 LOSNLNDIQALACLSALSMITERHGLKEPKRVEELCNKITSSLKDHQSGQALEPTESKV 600
DB 541 LOSNLNDIQALACLSALSMITERHGLKEPKRVEELCNKITSSLKDHQSGQALEPTESKV 600

QY 584 LGALVELRKICTGLQRIFYKLKLEDLVSPSPSIIDKFLDLPF 626
DB 584 LGALVELRKICTGLQRIFYKLKLEDLVSPSPSIIDKFLDLPF 626

```

```

DB 601 LRALVELRKICTGLQRIFYKLKLEDLVSPSPSIIDKFLDLPF 643
RESULT 3
NR43 RAT
ID NR43 RAT STANDARD; PRT; 628 AA.
AC P51179; Q63516; Q9OWQ3;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Orphan nuclear receptor NR4A3 (Nuclear hormone receptor NOR-1/NOR-2)
DE (Neuron-derived orphan receptor 1/2).
GN Name=Nr4a3; Synonyms=Nor1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=95110348; PubMed=7811288;
RA Ohkura N., Hijikuro M., Yamamoto A., Miki K.;
RT "Molecular cloning of a novel thyroid/steroid receptor superfamily
RL gene from cultured rat neuronal cells.";
RL Biochem. Biophys. Res. Commun. 205:1959-1965(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=9600221; PubMed=7556683; DOI=10.1016/0014-5793(95)00998-0;
RA Petropoulos I., Part D., Ochoa A., Zakin M.M., Lamae E.;
RT "NOR-2 (neuron-derived orphan receptor), a brain zinc finger protein,
RL is highly induced during liver regeneration.";
RL FEBS Lett. 372:273-278(1995).
RN [3]
RP SEQUENCE OF 318-345 FROM N.A.
RC TISSUE=Hippocampus;
RX PubMed=7914660;
RA Pena de Ortiz S., Cannon M.M., Jamieson G.A. Jr.;
RT "Expression of nuclear hormone receptors within the rat hippocampus:
RL identification of novel orphan receptors.";
RL Brain Res. Mol. Brain Res. 23:278-283(1994).
CC -!- FUNCTION: Binds to the B1a response-element.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=NOR-1;
CC IsoId=P51179-1; Sequences=Displayed;
CC Name=2; Synonyms=NOR-2;
CC IsoId=P51179-2; Sequences=VSP_010085, VSP_010086;
CC -!- TISSUE SPECIFICITY: Expressed at high levels in cultured apoptotic
CC neuronal cells and fetal brain, and at low level in adult brain.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR4
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D38530; BAA07535.1; -.
DR EMBL: X86003; CAA59993.1; -.
DR EMBL: L19343; AAB46395.1; -.
DR PIR: JC2493; JC2493.
DR PIR: S66671; S66671.
DR HSP: P43354; IOVL.
DR TRANSFAC: T02768; -.
DR TRANSFAC: T04752; -.
DR RGD: 61882; Nr4a3.
DR InterPro: IPR000536; Hrmn_recept_lig.
DR InterPro: IPR001723; Stdhrmn_receptor.

```


DR InterPro; IPR008946; Str ncl_receptor.
 DR InterPro; IPR001628; Znf C4steroid.
 DR Pfam; PF00104; Hormone recep; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STRDHORMONER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR ProDom; PD000035; Znf C4steroid; 1.
 DR SMART; SM00430; HOLI; 1.
 DR SMART; SM00399; Znf C4; 1.
 DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
 KW Alternative splicing; DNA-binding; Nuclear protein; Receptor;
 KW Transcription regulation; Zinc-finger.
 FT DNA_BIND 294 359 Nuclear receptor-type.
 FT ZN_FING 294 314 C4-type.
 FT ZN_FING 330 354 C4-type.
 FT DOMAIN 442 492 Ligand-binding (Potential).
 FT DOMAIN 99 112 Poly-His.
 FT DOMAIN 113 116 Poly-Gln.
 FT DOMAIN 286 289 Poly-Ser.
 FT VARSPLIC 421 430
 FT VARSPLIC 431 628 YCPTDQATAG -> VSSVNRFEWF (in isoform 2).
 FT VARSPLIC 431 628 Missing (in isoform 2).
 FT CONFLICT 13 13 P -> Q (in Ref. 2).
 FT CONFLICT 18 18 A -> P (in Ref. 2).
 SQ SEQUENCE 628 AA; 68564 MW; 9AC285D9A65226D9 CRC64;
 Query Match 92.1%; Score 3074; DB 1; Length 628;
 Best Local Similarity 92.2%; Pred. No. 6.9e-155;
 Matches 582; Conservative 12; Mismatches 29; Indels 8; Gaps 4;
 QY 1 MPCVQAQYSPSPGSSYAAQTSSEYTTIMMPDYTKLTMDLGSSTETATATTSPLSIST 60
 DB 1 MPCVQAQYSPSPGSSYAAQTSSEYTTIMMPDYTKLTMDLGSSTETATATTSPLSIST 60
 QY 61 FVEGYSNVELPSCVYQM-----QRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHH 116
 DB 61 FMEGYSSCELPSCLYQMPSPGPRPLIKVEGRGEGYHHHHHHHHHHHHHHHHHHHHHH 118
 QY 117 IPPASSPEVLPSTSMYFKQSPSPPTTPAPQAGALWDELSPACSIAPGGLDPP 176
 DB 119 IPPSGPEVLPSTSMYFKQSPSPPTTPPGPQAGALWDELSPACSIAPGGLDPP 178
 QY 177 MKAVTVP-AGARFPLFHFKSPSPHPAPSPAGSHLGYDPTAAALSLPLGAAAAGSQA 235
 DB 179 MKAVPMAAARFPF-PKPSPPHPAPSPAGSHLGYDPTAAALSLPLGAAAAGSQA 237
 QY 236 ALESHPYGLPLAKRAAPLAPPLGLTPSPATSSLLGESPSLPSPSSSSSGEGTCAVC 295
 DB 238 AALEGHPYGLPLAKRTATLTPPLGLTASPTASSLLGESPSLPSPSSSSSGEGTCAVC 297
 QY 296 GDNAACOHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVPDKRRNRRCYCRFOKCLSV 355
 DB 298 GDNAACOHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVPDKRRNRRCYCRFOKCLSV 357
 QY 356 GMVKEVVRTSLKGRRLPSPKPSLPQSPSPSPPICMNALVRALTDSTPRDL 415
 DB 358 GMVKEVVRTSLKGRRLPSPKPSLPQSPSPSPPICMNALVRALTDSTPRDL 417
 QY 416 YSRYPCTDQAAGTDAHVQOQYNNLLTASIDVSRSWAEKIPGFTDLPKEDQTLLESAPL 475
 DB 418 YSRYPCTDQAAGTDAHVQOQYNNLLTASIDVSRSWAEKIPGFTDLPKEDQTLLESAPL 477
 QY 476 ELFVLRLSRSRTAEDKFFVFCNGLVLRQCLRGFGWELDSKDFSLNQLSNLDIQAAL 535
 DB 478 ELFVLRLSRSRTAEDKFFVFCNGLVLRQCLRGFGWELDSKDFSLNQLSNLDIQAAL 537
 QY 536 CLSALSMITERHGLKEPKEVEELCNKITSSLDHQSQKQALPPTESKVLGALVELRKICT 595
 DB 538 CLSALSMITERHGLKEPKEVEELCNKITSSLDHQSQKQALPPTESKVLGALVELRKICT 597
 QY 596 LGLQRIFYKLKLEDLVSPSPSIDKFLDTLPF 626
 FT VARSPLIC 430 627 Missing (in isoform 2).

DB 598 QGLQRIFYKLKLEDLVSPSPSIDKFLDTLPF 628
 RESULT 4
 NR43_MOUSE
 ID NR43_MOUSE STANDARD; PRT; 627 AA.
 AC O9QZB6; O9QZB5;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Orphan nuclear receptor NR4A3 (Orphan nuclear receptor TEC)
 DE (Translocated in extraskelatal chondrosarcoma).
 GN Name=NR4A3; Synonyms=Tec;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=20163713; PubMed=10701778; DOI=10.1089/104454900314636;
 RA Maltais A., Labelle Y.;
 RT "Structure and expression of the mouse gene encoding the orphan
 RT nuclear receptor TEC";
 RL DNA Cell Biol. 19:121-130(2000).
 CC -!- FUNCTION: Binds to the B1A response-element (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=Long;
 CC IsoId=O9QZB6-1; Sequence=Displayed;
 CC Name=2; Synonyms=Short, TECdeltaC;
 CC IsoId=O9QZB6-2; Sequence=VSP 010083, VSP 010084;
 CC -!- TISSUE SPECIFICITY: Ubiquitous. Highest levels of expression in
 CC brain.
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR4
 subfamily.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; AF191211; AAF05622.1; -;
 CC EMBL; AF191212; AAF05623.1; -;
 CC HSSP; P43354; IOVL.
 CC MGD; MGI:1352457; Nr4a3.
 CC InterPro; IPR00536; Hrmn recept lig.
 CC InterPro; IPR001723; Strdhmn receptor.
 CC InterPro; IPR008946; Str ncl_receptor.
 CC InterPro; IPR001628; Znf C4steroid.
 CC Pfam; PF00104; Hormone_recep; 1.
 CC Pfam; PF00105; zf-C4; 1.
 CC PRINTS; PR00398; STRDHORMONER.
 CC PRINTS; PR00047; STROIDFINGER.
 CC ProDom; PD000035; Znf C4steroid; 1.
 CC SMART; SM00430; HOLI_1.
 CC SMART; SM00399; Znf C4; 1.
 CC PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
 KW Alternative splicing; DNA-binding; Nuclear protein; Receptor;
 KW Transcription regulation; Zinc-finger.
 FT DNA_BIND 293 358 Nuclear receptor-type.
 FT ZN_FING 293 313 C4-type.
 FT ZN_FING 329 353 C4-type.
 FT DOMAIN 441 491 Ligand-binding (Potential).
 FT DOMAIN 99 111 Poly-His.
 FT DOMAIN 112 115 Poly-Gln.
 FT DOMAIN 285 288 Poly-Ser.
 FT VARSPLIC 420 429 YCPTDQATAG -> VSSMNAFEPL (in isoform 2).
 FT VARSPLIC 430 627 /FTId=VSP 010083.
 FT VARSPLIC 430 627 Missing (in isoform 2).

```

FT  SQ  SEQUENCE  627 AA;  68454 MW;  2F49CB4C8DE315A8 CRC64;
/FTIG=VSP 010084.
Query Match      91.5%;  Score 3054.5;  DB 1;  Length 627;
Best Local Similarity 91.8%;  Pred. No. 7.4e-154;
Matches 579;  Conservative 12;  Mismatches 31;  Indels 9;  Gaps 4;

QY  1  MPCVQAQYSPSPGSSYAAQTYSSYTTTEIMNPDYTKLTMDLSTGTEITATATSLPSIST 60
DB  1  MPCVQAQYSPSPGSSYAAQTYSSYTTTEIMNPDYTKLTMDLSTGTEITATATSLPSIST 60

QY  61  FVEGYSSNYELKSPCYQWQ----QRPLIKVEGGRAPSYHHHHHHHHHHHHHHHHHHHH 116
DB  61  FVEGYSSNYELKSPCYQWQ----QRPLIKVEGGRAPSYHHHHHHHHHHHHHHHHHHHH 117

QY  117  IPPASSPEDEVLPSTSMYFKQSPSTPTTTPAPPOAGALWDEALPSAPGCIAPGLDPP 176
DB  118  IPPASSPEDEVLPSTSMYFKQSPSTPTTTPAPPOAGALWDEALPSAPGCIAPGLDPP 177

QY  177  MKAVPTV-AGARFPLFHFKPSPPHPPAPSPAGGHHLYGDPDTAAALSLPLGAAAGSQA 235
DB  178  MKAVPPMAAAAFPIF-FKPSPPHPPAPSPAGGHHLYGDPDTAAALSLPLGAAAGSQA 236

QY  236  AALESHPYGLPLAKRAAPLAFPLGLTPTASSLLGESPLSPSPSSSGEGTCAVC 295
DB  237  ALEGHYPYGLPLAKRTATLTFFPLGLTASPTASSLLGESPLSPSPNRRSSSGEGTCAVC 296

QY  296  GNAACOHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDPKRRNRQYCRFOKCLSV 355
DB  297  GNAACOHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDPKRRNRQYCRFOKCLSV 356

QY  356  GMVKEVVRTSLKGRGRPLSPKSPLOQEPSPSPSPPICMNALVRLDSTPRDL 415
DB  357  GMVKEVVRTSLKGRGRPLSPKSPLOQEPSPSPSPPICMNALVRLDSTPRDL 416

QY  416  YSKYPTDQAAAGTDAEHVQOQFNLLTASIDVSRSWAEKI PGFTDLPKEDQTLLESAPL 475
DB  417  YSKYPTDQATAGTDAEHVQOQFNLLTASIDVSRSWAEKI PGFTDLPKEDQTLLESAPL 476

QY  476  ELFVLRLSIRSTAEDKVFVNCGLVLRLQCLRGFGWELDSIKDFSLNLQSLNLDIQALA 535
DB  477  ELFVLRLSIRSTAEDKVFVNCGLVLRLQCLRGFGWELDSIKDFSLNLQSLNLDIQALA 536

QY  536  CLSALSMITERGLKPEKVEBELCNKITSSLDKHQSKGOALEPTSKVGLGALVELRKICT 595
DB  537  CLSALSMITERGLKPEKVEBELCNKITSSLDKHQSKGOALEPTSKVGLGALVELRKICT 596

QY  596  LGLQRFYKLJEDLVSPRSIIDKLFDTLPF 626
DB  597  LGLQRFYKLJEDLVSPRSIIDKLFDTLPF 627

RESULT 5
O97727  PRELIMINARY;  PRT;  446 AA.
AC  O97727;
DT  01-MAY-1999 (TrEMBLrel. 10, Created)
DT  01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Neuron-derived orphan receptor-1 beta.
GN  Name=NOR-1;
OS  Sus scrofa (Pig).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX  NCBI_TaxID=9823;
RN  [1]_
SEQUENCE FROM N.A.
RA  Cases-Langhoff C.; Castello A.; Martinez-Gonzalez J.; Badimon L.;
RL  Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC  -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC  -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR  EMBL; AJ011768; CAA03764.1; -.
HSSP; P43354; 10VL.

```

```

DR  GO; GO:0005634; C:nucleus; IEA.
DR  GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IEA.
DR  GO; GO:0003700; F:transcription factor activity; IEA.
DR  GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR  InterPro; IPR003072; NOR1_receptor.
DR  InterPro; IPR003070; Nucleoph receptor.
DR  InterPro; IPR001628; Znf_C4steroid.
DR  Pfam; PF00105; Zf-C4; 1_NUCLEAR.
DR  PRINTS; PRO1286; NORNUCCEPTR.
DR  PRINTS; PRO1284; NUCLEARECPT.
DR  PRINTS; PRO0047; STROIDFINGER.
DR  ProDom; PD000035; Znf_C4steroid.
DR  SMART; SM00399; Znf_C4; 1.
DR  PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW  DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
Transcription regulation; Zinc; Zinc-finger.
SQ  SEQUENCE  446 AA;  47716 MW;  E21960657606FB92 CRC64;

Query Match      62.9%;  Score 2099.5;  DB 2;  Length 446;
Best Local Similarity 89.9%;  Pred. No. 1.8e-103;
Matches 391;  Conservative 9;  Mismatches 18;  Indels 17;  Gaps 3;

QY  1  MPCVQAQYSPSPGSSYAAQTYSSYTTTEIMNPDYTKLTMDLSTGTEITATATSLPSIST 60
DB  1  MPCVQAQYSPSPGSSYAAQTYSSYTTTEIMNPDYTKLTMDLSTGTEITATATSLPSIST 60

QY  61  FVEGYSSNYELKSPCYQWQ----RPLIKVEGGRAPSYHHHHHHHHHHHHHHHHHHHH 110
DB  61  FVEGYSSNYELKSPCYQWQ----RPLIKVEGGRAPSYHHHHHHHHHHHHHHHHHHHH 120

QY  111  --HQOQPSIPPASSPEDEVLPSTSMYFKQSPSTPTTTPAPPOAGALWDEALPSAPGCTA 168
DB  121  PPQOQPSIPPASSPEDEVLPSTSMYFKQSPSTPTTTPAPPOAGALWDEALPSAPGCTA 180

QY  169  PGFLDLPDMKAVPTVAGARFPLFHFKPSPPHPPAPSPAGGHHLYGDPDTAAALSLPLG-- 226
DB  181  PGFLDLPDMKAVPTVAGARFPLFHFKPSPPHPPAPSPAGGHHLYGDPDTAAALSLPLGAA 240

QY  227  ---AAAAAGSQAALLESHPYGLPLAKRAAPLAFPLGLTPTASSLLGESPLSPSPSR 283
DB  241  AAAAAAGSQAALLESHPYGLPLAKRAAALAFSPGLTPTASSLLGESPLSPSPNR 300

QY  284  SSSSGEGTCAVCGDNAACHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDPKRRNR 343
DB  301  STASGEGTCAVCGDNAACHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDPKRRNR 360

QY  344  COYCFQKCLSVGMVKEVVRTSLKGRGRPLSPKSPLOQEPSPSPSPPICMNALV 403
DB  361  COYCFQKCLSVGMVKEVVRTSLKGRGRPLSPKSPLOQEPSPSPSPPICMNALV 420

QY  404  RALTDSTPRDLDSYR 418
DB  421  RALTDSTPRDLDSYR 435

RESULT 6
NR422 HUMAN
ID  NR422 HUMAN  STANDARD;  PRT;  598 AA.
AC  P43354; Q16311;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  25-OCT-2004 (Rel. 45, Last annotation update)
DE  Orphan nuclear receptor NR4A2 (Orphan nuclear receptor NURR1)
DE  (Immediate-early response protein NOT) (Transcriptionally inducible
nuclear receptor).
GN  Name=NR4A2; Synonyms=NOT, NURR1, TINUR;
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]_
SEQUENCE FROM N.A.
RP  TISSUE=Blood;

```



```

QY 444 SIDVSRWAKEIPGFTDLPKEDQTLIESAFLEFVLRSLRSNTAEDKFVFCNGVLVLR 503
DB 413 SMEIIRGWAKEIPGFADLPKADODLLFESAFLEFVLRSLRSNPVPEGKLIFCNGVVLHR 472
QY 504 LCLRGFGWLSIDISDFSLNLSQSLNDIOALACLALSAMITRHHGLKEPKRVEELCNKIT 563
DB 473 LQCVRGFGWLSIDISDFSLNLSQSLNDIOALACLALSAMITRHHGLKEPKRVEELCNKIT 532
QY 564 SSKLDHQ--SKQOALEPTE-SKVLGALVELRKICTLGLQRIPLYLKLEDLVSPSIDKLF 620
DB 533 NCLKDHVTNNGLRPNVLSKLLGLKPELRLCTQGLQRIPLYLKLEDLVPPPAIDKLF 592
QY 621 LDTLPF 626
DB 593 LDTLPF 598

RESULT 7
NR42_MOUSE
ID NR42_MOUSE STANDARD; PRT; 598 AA.
AC Q06219; O08690;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Orphan nuclear receptor NR4A2 (Orphan nuclear receptor NURR1) (NUR-
  related factor 1).
GN Name=NR4A2; Synonyms=Nurr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93149122; PubMed=1491694; DOI=10.1210/me.6.12.2129;
RA Law S.W.; Connely O.M.; Denayo P.J.; O'Malley B.W.;
RT "Identification of a new brain-specific transcription factor, NURR1."
RL Mol. Endocrinol. 6:2129-2135(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=97288523; PubMed=9143501; DOI=10.1006/geno.1997.4677;
RA Castillo S.O.; Xiao Q.; Lyu M.S.; Kozak C.A.; Nikodem V.M.;
RT "Organization, sequence, chromosomal localization, and promoter
  identification of the mouse orphan nuclear receptor Nurrl gene."
RL Genomics 41:250-257(1997).
CC -!- FUNCTION: Probable nuclear receptor that regulates gene expression
  in brain tissue.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
  Event=Alternative splicing; Named isoforms=2;
  Name=NURR1;
  IsoId=Q06219-1; Sequence=Displayed;
  Name=NURRIA;
  IsoId=Q06219-2; Sequence=VSP_003710, VSP_003711;
CC -!- TISSUE SPECIFICITY: Brain.
CC -!- DEVELOPMENTAL STAGE: Expression begins in the embryo, increases in
  neonates and decreases in the adult.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR4
  subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; S53744; AAB25138.1; -
DR EMBL; U86783; AAC53153.1; -
DR EMBL; U86783; AAC53154.1; -

```

```

DR PIR; A46225; A46225.
DR HSSP; P43354; 1OVL.
DR TRANSPAC; T04312; -.
DR MGD; MGI:1352456; Nr4a2.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0007399; P:neurogenesis; IMP.
DR GO; GO:0042053; P:regulation of dopamine metabolism; IDA.
DR InterPro; IPR000536; Hrmn_recept_lig.
DR InterPro; IPR001723; Stchrnm_receptor.
DR InterPro; IPR008946; Str_ncl_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; Zf-C4; 1.
DR PRINTS; PR00398; STROHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Alternative splicing; DNA-binding; Nuclear protein; Receptor;
KW Transcription regulation; Zinc-finger.
FT DOMAIN 80 138
FT DOMAIN 127 233
FT DNA_BIND 263 328
FT ZN_FING 263 283
FT ZN_FING 299 323
FT DOMAIN 346 391
FT DOMAIN 409 459
FT VARSPLIC 455 455
FT VARSPLIC 456 598
FT Missing (in isoform NURR1A).
FT /FTId=VSP_003710.
FT /FTId=VSP_003711.
SQ SEQUENCE 598 AA; 66592 MW; 66592 MW; DIACSEB2FFB8D366 CRC64;

Query Match 52.3%; Score 1746; DB 1; Length 598;
Best Local Similarity 54.4%; Pred. No. 1.2e-84;
Matches 362; Conservative 75; Mismatches 121; Indels 108; Gaps 18;

QY 1 MPCVQAOVSPSPGGSYAAQTY-----SSSYTEINPNPDYTKLMDLGSTEITATTSLP 56
DB 1 MPCVQAOVSPSPGGSYAAQTY-----SSSYTEINPNPDYTKLMDLGSTEITATTSLP 58
QY 57 STSTVEGVSSNYELKPSCVVQM-----GRPLKVEGRAPSVHHHHHHHHHHHHHHHQQQH 112
DB 59 SPSTFMDNYSTGYDVKPCLYQMPLSGQSSIKVEDIQHNTYQOSH----- 105
QY 113 QQPSIPPASSPEDEVLP-STSMYFKQSPSPSTTPAFPQQAGALWDEALPSAPGCIAPGP 171
DB 106 ---LPPQS---EEMPHSGSVYKPSPPPTSTFQVQHSPPMDD-----PGS 148
QY 172 LLDPPMKAVPT-----VAGARPLFHFKPSPPHPPA-----PSPAG 207
DB 149 LHNPHQNYVATHTMIEQRKTPVSRSLSLFSFKQSPGTPVSSCCMRFDGLPHVPMNPEPAG 208
QY 208 GHHLGYDPTAAALSLPLGAAAAGSQAALSHSPVGLP-LAKRAAPLAFPLGLTPSPPT 266
DB 209 SHHV-----VDQGFAPVNPTRKPSMGFFGLQI---GH 239
QY 267 ASSLLGESPSLPSPSSSSGEGTCAVCGDNAAACQHVGVRTCEGCKGFFKRTVQKNKY 326
DB 240 ASQLL--DTQVPSPPSRGSPSNEGLCAVCGDNAAACQHVGVRTCEGCKGFFKRTVQKNKY 297
QY 327 VCLANKNCVDRKRRNRNRCQYCFQKCLSVGMVKEVVRTDLSKGRGRRLPSKPSKPIQQBP 386
DB 298 VCLANKNCVDRKRRNRNRCQYCFQKCLAVGMVKEVVRTDLSKGRGRRLPSKPSKPIQQBP 352
QY 387 SOPSPPSPICMNNALVRALTDSTP--RDLDYSRY-CPTDQAAAGTDASHVOQFYNLLTA 443
DB 353 QDPSPPSPVSLISALVRAHVDNSNPAMTSLDYSRFQANPDYQMSGDDTQHIQQFYDLLTG 412
QY 444 SIDVSRWAKEIPGFTDLPKEDQTLIESAFLEFVLRSLRSNTAEDKFVFCNGVLVLR 503
DB 413 SMEIIRGWAKEIPGFADLPKADODLLFESAFLEFVLRSLRSNPVPEGKLIFCNGVVLHR 472

```

QY 504 LQCLRGFGWLSIKDFSLNLSQSLNDIQAALCLSLSMITRHHGLKEPKRVEELQNKIT 563
 DB 473 LQCVRGFGWLSIVFSSNLQNNIDISAFSCIAALAMVTRHGLKEPKRVEELQNKIV 532
 QY 564 SSIKQHQ--SKQALPTE-SKVLGALVELKICTLGLORIFVYKLEDLVSPSIIKLF 620
 DB 533 NCLKHVHTVNGLNRPNYLSKLLGLPELRLTCLQGLORIFVYKLEDLVPPPAIDKLF 592
 QY 621 LDTLPF 626
 DB 593 LDTLPF 598
 RESULT 8
 NR42 RAT
 ID NR42 RAT STANDARD; PRT; 598 AA.
 AC Q07917; O35865;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Orphan nuclear receptor NR4A2 (Orphan nuclear receptor NURR1) (NUR-
 DE related factor 1) (Regenerating liver nuclear receptor 1) (RNR-1) (SL-
 DE 222) (Nuclear orphan receptor HZF-3).
 GN Names=NR4A2; Synonyms=Hzi-3, Nurr1, Rnr1;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Hippocampus;
 RX MEDLINE=93232045; PubMed=7914660; DOI=10.1016/0169-328X(94)90235-6;
 RA de Ortiz S., Cannon M.M., Jamieson G.A.;
 RT "Expression of nuclear hormone receptors within the rat hippocampus:
 RT identification of novel orphan receptors.";
 RL Brain Res. Mol. Brain Res. 23:278-283(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93232045; PubMed=8473329;
 RA Scearce L.M., Laz T.M., Hazel T.G., Lau L.F., Taub R.;
 RT "NR-1, a nuclear receptor in the NGFI-B/Nur77 family that is rapidly
 RT induced in regenerating liver.";
 RL J. Biol. Chem. 268:8855-8861(1993).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=97364974; PubMed=9221923; DOI=10.1016/S0169-328X(97)00056-9;
 RA Xing G.O., Zhang L.X., Zhang L., Heynen T., Li X.L., Smith M.A.,
 RA Weiss S.R.B., Feldman A.N., Detera-Wadleigh S., Chuang D.M.,
 RA Post R.M.;
 RT "Rat nurrl is prominently expressed in perirhinal cortex, and
 RT differentially induced in the hippocampal dentate gyrus by
 RT electroconvulsive vs. kindled seizures.";
 RL Brain Res. Mol. Brain Res. 47:251-261(1997).
 CC -!- FUNCTION: Probable nuclear receptor. May confer liver-specific
 CC regulation of delayed-early genes induced later in the G1 phase of
 CC regeneration along with HMR.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Not expressed in quiescent liver but is
 CC rapidly induced following partial hepatectomy and is specific to
 CC hepatic growth as it is not induced in other mitogen-treated
 CC cells. Expressed at very low levels in the lung, spleen and
 CC stomach and at high levels in the brain.
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR4
 CC subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U01146; AAC52143.1; -
 DR EMBL; L08595; AAC42058.1; -
 DR EMBL; U72345; AAC53315.1; -
 DR PIR; I84692; I84692.
 DR HSSP; P43354; LOVL.
 DR TRANSFAC; T04743; -
 DR RGD; 3202; Nr4a2.
 DR InterPro; IPR000536; Hrmn_recept_lig.
 DR InterPro; IPR001723; Strhmn_receptor.
 DR InterPro; IPR008946; Strncl_receptor.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00104; Hormone_recep; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STRDHORMONER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOLI; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
 KW Zinc-finger.
 FT DOMAIN 80 138 Gln-rich.
 FT DOMAIN 127 233 Pro-rich.
 FT DNA_BIND 263 328 Nuclear receptor-type.
 FT ZN_FING 263 283 C4-type.
 FT ZN_FING 299 323 C4-type.
 FT DOMAIN 346 391 Pro-rich.
 FT DOMAIN 409 459 Ligand-binding (Potential).
 FT CONFLICT 182 182 R -> A (in Ref. 2).
 FT CONFLICT 182 182 R -> P (in Ref. 3).
 FT CONFLICT 250 250 S -> P (in Ref. 2).
 FT CONFLICT 252 252 P -> A (in Ref. 1).
 FT CONFLICT 362 362 V -> G (in Ref. 2).
 FT CONFLICT 364 371 LISALVRA -> DQCPROT (in Ref. 2).
 FT CONFLICT 434 434 D -> S (in Ref. 2).
 SQ SEQUENCE 598 AA; 66621 MW; 0A54ED59FB70DEA8 CRC64;
 Query Match 52.0%; Score 1736; DB 1; Length 598;
 Best Local Similarity 54.2%; Pred. No. 4.2e-84;
 Matches 361; Conservative 74; Mismatches 123; Indels 108; Gaps 18;
 QY 1 MPCVQAYSPSPGSSYAAQTY----SSEYTEINPDYTKLTMDLSTETITATTSILP 56
 DB 1 MPCVQAYGSSPQSGPASQSYSHSGEYSSDLTFEFVKESMDLTWEI--TATTSILP 58
 QY 57 SISTVEGYSNYELKPSQVYQ-----ORPLIKVEGRAPSYHHHHHHHHHHHHHHHHH 112
 DB 59 SFSTFMDNYSTGYDVKKPCLYQMLPSGQSSIKVEDIQMHNYQOHSH----- 105
 QY 113 QQPSTPPRASSPEDEVLIP-STSMYFKQSPSPSTPTTAPFPPOAGALWDEALPSAPGCIACP 171
 DB 106 ----LPPQS-----EEMPHSGSVYKPSPPPTSTPGFQVQSHSPWMD-----PQS 148
 QY 172 LLDPEMKAVPT-----VAGARFPLFHFKPSPHPPEA-----PSPAG 207
 DB 149 LHNHFQNYVATHMTIEQRKTPVSRLSLSFSKQSRGTPVSSCQMFQDGLHVPMPPEPAG 208
 QY 208 GHHLGYDPTAAALSLPIGAAAAAGSQAALSHHYGLP-LAKRAAPLAPPLGLTTPSP 266
 DB 209 SHHV-----VDGQTEAVENPIRKPSMGFPGQLI---GH 239
 QY 267 ASSLLGESPSLPSPSSSSSGEGTCVCGDNAAOCHYGVRTCEGCKGFKRTVQKNKY 326
 DB 240 ASQLL--DTQVPSPPSPSGPSNEGCAVCGDNAAOCHYGVRTCEGCKGFKRTVQKNKY 297
 QY 327 VCLANKNCVDPKRRNRQCYCRFQKLSVGMVKEVVRTDSLKGRRLSPKSPKSLQOEP 386
 DB 298 VCLANKNCVDPKRRNRQCYCRFQKLSVGMVKEVVRTDSLKGRRLSPKSPKSPKSPK 352
 QY 387 SQSPSPSPPICMNAALVRLTDTSTP--RDLSRY-CFTDQAAAGTDAHVQOQFNLLTA 443

```

Db 353 QDPSPSPVSLISALVRAHVSNAFTSLDSRFOANPDYQMGDDTHQIOQFDLLTG 412
Qy 444 SIDVRSWAKEIPGFTDLPKEDOTLLIESAFLEFVLRLSIRNTAEDKFCVFCNGLVLR 503
Db 413 SMEIIRGWAKEIPGADLPKADODLLFESAFLEFVLRLAYRSNPVEGKLIFCNGWLHR 472
Qy 504 LOCLRGFGWELSDIKDFSLNLSQSLNLDIOALACLNSLMTIRHGLKEPKRVBELCNKIT 563
Db 473 LQCVRGFGWELSDIVFESNLQNMNIDISAFSCIAALAMVTERHGLKEPKRVBELCNKIT 532
Qy 564 SSKDKHO--SKGQALEPTE-SKVLGALVELKICTGLQRIEVLKLEDLVSPSIIDKLF 620
Db 533 NCLKHVTFNNGLNRPVYLSKLGKLPRLICTQGLQRIEVLKLEDLVPPPAIDKLF 592
Qy 621 LDTLPP 626
Db 593 LDTLPP 598

RESULT 9
Q98TQ3
ID Q98TQ3 PRELIMINARY; PRT; 600 AA.
AC Q98TQ3;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Nuclear receptor related 1.
GN Name=nurrl;
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kapsimali M., Bourrat F., Vernier P.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; AJ278700; CAC27783.1; -.
DR HSP; P43354; 10VL.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000536; Hrmn_recept_lig.
DR InterPro; IPR003070; Nucorpb_receptor.
DR InterPro; IPR003073; NURR_receptor.
DR InterPro; IPR001723; Strdhmn_receptor.
DR InterPro; IPR008946; Str_ncl_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR01284; NUCLEARRECPTR.
DR PRINTS; PR01287; NURRNUCRCPTR.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI_1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 600 AA; 66925 MW; BBSPE169B7DBFBAD CRC64;

Query Match 50.3%; Score 1677; DB 2; Length 600;
Best Local Similarity 53.7%; Pred. NO. 5.5e-81;
Matches 353; Conservative 81; Mismatches 135; Indels 88; Gaps 19;

Qy 1 MPCVQAQYSPSGSYAQTYS-----SEYTTIEMNPDTYKLTMDLGSFEITATATSLP 56

```

```

Db 1 MPCVQAQYSGSPASQSYHYTAGESCDLTPFVKFMDLTNTETI--TATSLP 58
Qy 57 S1STFVEGYSNYLKPSCVYQM-----QRPLIKVEGRAPSYHHHHHHHHHHHHHHHHH 112
Db 59 SFSTMDNYNTGVDKPPCLYQMPHSGEQSIKVEDVQMHSY-----H 101
Qy 113 QOPSTPRASSPEDEVLPSTSMYFKOSPSTPTTPAPPOAGALMDEALPSAPCCI----- 167
Db 102 QOOSHLPPQS--EEMIAHTGPMYFKPSSPHAPSTPNFQVQPNHMD-----PGSLHSFHQ 154
Qy 168 ----APGPLLDPMKAVPTVAGARPLFHFKPSPPHPPAPSPAGGHHGLGYDPTAAALSL 223
Db 155 NYVAATSHWE--QRKTPV---SRLSLFSFKQSPGCTPVSS---COMFEDG----- 197
Qy 224 PLGAAAAAGSQA--ALESHPYGLPLA-KRAAPLAPPLPLGLTPSTASSLLGESPEL----- 277
Db 198 PLHVSMTNDNPGAHRLDQSPAVPSAIRKQAGLAPP-----HSLQLSHGHQLVDS 248
Qy 278 --PSPPSSSSSGEGTCVACGDNAACOHYGVRTCEGCKGFFKRTVQKNKYYCLANKNCP 335
Db 249 QVPSPSPSGSPSNEGLCAVCGDNAACOHYGVRTCEGCKGFFKRTVQKNKYYCLANKNCP 308
Qy 336 VDKRRNRRCQYCRFOKCLSVGMVKEVVRTDSLKGRRGLRPSKPSKPSKPSKPSPPSP 395
Db 309 VDKRRNRRCQYCRFOKCLSVGMVKEVVRTDNLKGRRLPSKPSKPSKPSPPSP 363
Qy 396 ICMNVALVALTDTSP--RDLDSRY-CPTDQAAAGTAEHVQVFNLTASIDVRSWA 452
Db 364 VSLISALVRAHVDSNPSMADYSRFOANPDYQMTQDNTQHTIQFYDILLTGSMEIIRGA 423
Qy 453 EKIPGFTDLPKEDOTLLIESAFLEFVLRLSIRNTAEDKFCVFCNGLVLRLOCLRGFGE 512
Db 424 EKIPGFTDLPKEDOTLLIESAFLEFVLRLSIRNTAEDKFCVFCNGLVLRLOCLRGFGE 483
Qy 513 WLDSIKDFSLNLSQSLNLDIOALACLNSLMTIRHGLKEPKRVBELCNKITSLKDH--- 569
Db 484 VDAIVEFSSNLQSLDIDISAFSCIAALAMVTERHGLKEPKRVBELCNKITSLKDH--- 543
Qy 570 QSKGQALEPTESKVLGALVELKICTGLQRIEVLKLEDLVSPSIIDKLFDTLPP 626
Db 544 NSGGLNRPVYLSKLGKLPRLICTQGLQRIEVLKLEDLVPPPAIDKLFDTLPP 600

RESULT 10
Q98TQ3
ID Q98TQ3 PRELIMINARY; PRT; 586 AA.
AC Q6DH08;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Zgc:92696.
GN Name=zgc:92696;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max J., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stabenstein M.J., Udwin T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

```


RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grinnwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalala U., Smailus D.E., Schnerch A., Schein J.E.,
RA	Jones S.J., Marra M.A.;
ET	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Eye;
RA	Straussberg R.;
RL	Submitted (JUL-2004) to the ENBL/GenBank/DBSJ databases.
CC	-I- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC	-I- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR	EMBL; BC076176; AAH76176.1; -
DR	GO; GO:0005634; C:nucleus; IEA.
DR	GO; GO:0003707; F:steroid hormone receptor activity; IEA.
DR	GO; GO:0003700; P:transcription factor activity; IEA.
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR	InterPro; IPR000536; Hmon_recept_lig.
DR	InterPro; IPR003070; Nucoroph_receptor.
DR	InterPro; IPR003073; NURR_receptor.
DR	InterPro; IPR001723; Stchrnm_receptor.
DR	InterPro; IPR008946; Str_ncl_receptor.
DR	InterPro; IPR001628; Znf_Casteroid.
DR	Pfam; PF00104; Hormone_recep; 1.
DR	Pfam; PF00105; zf-C4; 1.
DR	PRINTS; PR01284; NUCLEARRECPTR.
DR	PRINTS; PR01287; NURRNURCPTFR.
DR	PRINTS; PR00398; STRDHORMONER.
DR	PRINTS; PR00047; STROIDFINGER.
DR	ProDom; PD000035; Znf_Casteroid; 1.
DR	SMART; SM00430; HOL1; 1.
DR	SMART; SM00399; Znf_C4; 1.
DR	PROSITE; PS00033; NUCLEAR_RECEPCTOR; 1.
KW	DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW	Transcription regulation; Zinc; Zinc-finger.
SQ	SEQUENCE 586 AA; 65343 MW; 6A91CD8BC0316769 CRC64;
	Query Match 47.6%; Score 1588; DB 2; Length 586;
	Best Local Similarity 53.7%; Pred. No. 2.8e-76;
	Matches 347; Conservative 74; Mismatches 145; Indels 80; Gaps 19;
Qy	1 MPCVQAQYSPSPGSSVAATYS----SEYTEIMNPDYKLTMDLGSSTEITATATSLP 56
Dd	: : : : : : : :
Dd	1 MPCVQAQYSPSPGASPASQSYSNTTGEYNCDFLTFFVKFSMDLNAEI--AVTSLLP 58
Qy	57 SISTFVEGYSNELKPSCVQM----QRPLIKVEGRAPSYHHHHHHHHHHHQOQH 112
Dd	: : : : : : : :
Dd	59 SFSTFDVTSSNYDVKKPCLYQMAHSGDQLSIKVEEIPA-----HCYHQOQH 105
Qy	113 QOPSIPASSPEDVLST-SMYFKQSPPTTPPPAQAGALWDE--ALPS-APGCIA 168
Dd	: : : : : : : :
Dd	106 HQP-----HQAEESIPHTGAIIYKPSNPISQSPSYPTAPHHTWTEDTSLHSFQNYLA 159
Qy	169 PGPLLDPMKAVPTVAGARRPLRFHKSPHPHPAPSAPA---GGHLGLYDPTAAALSLSLP 224
Dd	: : : : : : : :
Dd	160 TSHIMDQRK-----NAMSLRFKQSPVDVTPMPSQMPFGDSLHVSMGPDTH----- 207
Qy	225 LGAAAAAGSQAALESHPYGLPLAKRAAPLAFPLGLTPTSPATSSLGESPSLSPSPSR 284
Dd	: : : : : : : :
Dd	208 -----HRALDS--FALP-----APPRKHQGVGLSHSLNVGHPLLESP-VASPOARG 250
Qy	285 SSSEGTCACVGDNAAACHGVGVRTCEGCKGFKKTVOKNAKYVCIANKNCPVDKRRNR 344
Dd	: : : : : : : :
Dd	251 SPSEGHLCACVGDNAAACHGVGVRTCEGCKGFKKTVOKNAKYVCIANKNCPVDKRRNR 310
Qy	345 QYCRFOKLCSVMGWKEVVVRDLSKGRGRPLSPKPKSPLQGEPSOPSPS-PPICMMNALV 403
Dd	: : : : : : : :
Dd	311 QYCRFOKLCSVMGWKEVVVRASLKGRGRPLSPKPKSP-----QDIPIVFTPVNLINALV 364
Qy	404 RALTDTSP--RDLDYSRYCTDQAAAGTDAAE-HVQQFYNNLTASTIDSYSRWAEKIPGFTD 460


```

Db      217 QUGER-----ESY-----SISTAPGL-----APTSPHLDQPGMLDAP-VPSAKARSG 258
QY      286 --SSGEGTCAVCGDNAAACHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVPDKRRNR 343
Db      259 APSGSEGRCAVCGDNAAACHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVPDKRRNR 318
QY      344 CQYCRFQKCLSGVMKVEVVRTDSLKGRRGLSPKPSPLQOEPSOPSPSPPCIMNALV 403
Db      319 CQCFRQKCLAVGMKVEVVRTDSLKGRRGLSPKQP-----PDASPANLLTSILV 369
QY      404 RALTDSTP--RDLDSRYCPTDQAAAG--TDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTD 460
Db      370 RAHLDSPGSTAKLDYSKFQELVLPHFHGKEDAGDVQOQFYDLLSGSLEIVIRKWAEEKIPGFAE 429
QY      461 LPKEDOTLLIESAFLEFVLRLSIRSNATAEKDFVFCNGLVLRHLOCLRGFGWLDISIKDF 520
Db      430 LSPGQDILLIESAFLEFVLRLSIRSNATAEKDFVFCNGLVLRHLOCLRGFGWLDISILAF 489
QY      521 SINLOSINLDIOALACLSALSMITERHGLKEPKRVEELCNKITSSIKDHQSKGOALEPTE 580
Db      490 SRSLSGLVVDPAFACLSALVLTDRHGLQEPFRVEELQNRIASCLKEHVS-AVAGEPOP 548
QY      581 ----SKVLGALVELRKICTLGLQRIFYLYKLEDLVSPPSIIDKLFDTLPP 626
Db      549 ASCLSRLLGKLPRLTCTQGLQRIFYLYKLEDLVSPPSIIDKLFMDTLPP 598

RESULT 13
Q8N3V2
ID      Q8N3V2      PRELIMINARY;      PRT;      598 AA.
AC      Q8N3V2;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE      Hypothetical protein DKFZp451G125.
GN      Name=DKFZp451G125;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]_TaxID=9606;
RP      SEQUENCE FROM N.A.
RC      TISSUE=Skeletal muscle;
RG      The German cDNA Consortium;
RA      Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA      Fobo G., Han M., Wiemann S.;
RL      Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC      -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC      -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR      EMBL; AL831844; CAD38550.2; -.
DR      GO; GO:0005634; C:nucleus; IEA.
DR      GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IEA.
DR      GO; GO:0003700; F:transcription factor activity; IEA.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
KW      DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
KW      Receptor; Transcription; Transcription regulation; Zinc; zinc-finger.
SQ      SEQUENCE 598 AA; 64445 MW; C96154E06013AC8B CRC64;

Query Match          41.0%; Score 1368; DB 2; Length 598;
Best Local Similarity 48.5%; Pred. NO. 1.2e-64;
Matches 315; Conservative 80; Mismatches 179; Indels 76; Gaps 21;

QY      1 MPCVQAQY---SPSPGGSSAAQYTSSEYTTIMNPDYTKLTMDLGSTEITATATSLPS 57
Db      1 MPCVQAQYTPAPSPGPRDHLA-----SDPLTPERIKPTMDLASPEAAPATAUPS 52
QY      58 ISTFVEGYSNVELKPCVYQMQ---RPLKVEGRAPSYHHHHHHHHHHHHHHHHQOQ 114
Db      53 FTTFMDYGTGEFD---TFYQLPGTVQPCSSASSASSSTSSSATSPASAFKFDQVY 109
QY      115 PSIP-PASSPEPEVLPTSMYFKQGPS--TPTTAPFPQACALWDEAL-PSAPGCIAPG 170

```

```

Db      110 GCYCPGRSGPVDEALSSSGSDYGYGSPCAPSPSTPFPQPLSPWDGFGHSSPSQTYEG 169
QY      171 PLLDDPMKAVPTVAGARFPLFHFKSPSPHPAPSPAGGHHLGYDPTAAALSL-PLGAAA 229
Db      170 --LRATWQLPKASGPPQPAFFSPPTGSPS-----LAQSPKLFPQATH 216
QY      230 AAGSOAALESHPYGLPLAKRAAPLAPPLGLTPSPSTASSLLG-----ESPSPSPSPSRSS 285
Db      217 QLQ-----EGESYSM-----PTAAPPGL-----APTSPHLESGILDTP-VTSTKARSG 258
QY      286 SSG--EGTCVAVCGDNAAACHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVPDKRRNR 343
Db      259 APGSGEGRCAVCGDNAAACHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVPDKRRNR 318
QY      344 CQYCRFQKCLSGVMKVEVVRTDSLKGRRGLSPKPSPLQOEPSOPSPSPPCIMNALV 403
Db      319 CQCFRQKCLAVGMKVEVVRTDSLKGRRGLSPKQP-----PDASPANLLTSILV 369
QY      404 RALTDSTP--RDLDSRYCPTDQAAAG--TDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTD 460
Db      370 RAHLDSPGSTAKLDYSKFQELVLPHFHGKEDAGDVQOQFYDLLSGSLEIVIRKWAEEKIPGFAE 429
QY      461 LPKEDOTLLIESAFLEFVLRLSIRSNATAEKDFVFCNGLVLRHLOCLRGFGWLDISIKDF 520
Db      430 LSPGQDILLIESAFLEFVLRLSIRSNATAEKDFVFCNGLVLRHLOCLRGFGWLDISILAF 489
QY      521 SINLOSINLDIOALACLSALSMITERHGLKEPKRVEELCNKITSSIKDHQSKGOALEPTE 580
Db      490 SRSLSGLVVDPAFACLSALVLTDRHGLQEPFRVEELQNRIASCLKEHVA-AVAGEPOP 548
QY      581 ----SKVLGALVELRKICTLGLQRIFYLYKLEDLVSPPSIIDKLFDTLPP 626
Db      549 ASCLSRLLGKLPRLTCTQGLQRIFYLYKLEDLVSPPSIIDKLFMDTLPP 598

RESULT 14
NR41_HUMAN
ID      NR41_HUMAN      STANDARD;      PRT;      598 AA.
AC      P22736;
DT      01-AUG-1991 (Rel. 19, Created)
DT      01-AUG-1991 (Rel. 19, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Orphan nuclear receptor NR4A1 (Orphan nuclear receptor HMR) (Early
DE      response protein NAK1) (TR3 orphan nuclear receptor).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Fetal skeletal muscle;
RC      MEDLINE=91133413; PubMed=2283997;
RA      Nakai A., Kartha S., Sakurai A., Toback F.G., Degroot L.J.;
RT      "A human early response gene homologous to murine nur77 and rat NGFI-
RL      B, and related to the nuclear receptor superfamily.";
RL      Mol. Endocrinol. 4:1438-1443(1990).
RN      [2]
RP      SEQUENCE FROM N.A.
EX      MEDLINE=90173205; PubMed=2626032; DOI=10.1016/0022-4731(89)90114-3;
RA      Chang C., Kokontis J., Liao S., Chang Y.;
RT      "Isolation and characterization of human TR3 receptor: a member of
RT      steroid receptor superfamily.";
RL      J. Steroid Biochem. 34:391-395(1989).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Skin;
RX      MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```


DR	GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR	InterPro: IPR003071; HMR_receptor.
DR	InterPro: IPR000536; Hrmn_recept_lig.
DR	InterPro: IPR003070; Nucoriph_receptor.
DR	InterPro: IPR001723; Stdhrmn_receptor.
DR	InterPro: IPR008946; Str_ncl_receptor.
DR	InterPro: IPR001628; Znf_Casteroid.
DR	Pfam: PF00104; Hormone_recep; 1.
DR	Pfam: PF00105; zf-C4; 1.
DR	PRINTS: PR01285; HMNNUCREPTR.
DR	PRINTS: PR01284; NUCLEARRECPTR.
DR	PRINTS: PR00398; STRDHORMONER.
DR	PRINTS: PR00047; STROIDFINGER.
DR	ProDom: PD000035; Znf_Casteroid; 1.
DR	SMART: SM00430; HOLI; 1.
DR	SMART: SM00399; ZNF C4; 1.
DR	PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW	DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW	Transcription regulation; Zinc; Zinc-finger.
SQ	SEQUENCE 598 AA; 64523 MW; 5B700BAVC25FDE33 CRC64;

Query Match		40.9%; Score 1364; DB 2; Length 598;
Best Local Similarity		47.5%; Pred. No. 2e-64;
Matches 317; Conservative		74; Mismatches 165; Indels 112; Gaps 23;

QY	1	MPCVQAQY----	SPSPGSSVAAYQSYSEYTEINMPDYTKLTMDLGSTETATATTSLPS	57
DB	1	: :	: :: :: ::	:
QY	58	ISTFVEGYSNVELKPSCVVQM-Q--	RPLIKVEGRAPSVHHHHHHHHHHHQQOHHQ	114
DB	53	FSTFMGTGTEFD---	TFLVLPGTVPCCSSASSSSATSPASAFKPEDFOVY	109
QY	115	PSIP-PASSPEDEVLPSTMYFKQSPPS--	TPTTPAF-PPAQ-----	153
DB	110	GCVPGPLSGPVDEALSSGSDDYVGSPCASPSPTSPFPQLSPWDGSGFHFSFSOTYEG	169	
QY	154	-ALWDEALPSAPGCIAEPLLDPPMKAVPTVAGARFPLFHFKSPSPHPAPPSPAGGHHLG	212	
DB	170	LRAWTEQLPKASG-----	PFQ-----PPAFFSFSPPTGSPSPS-----	201
QY	213	YDPTAAALSL-PLGAAAAAGSAQALESHPYGLPLAKRAAPLAFPPLGLTPSPASSLL	271	
DB	202	--LAQSPLKLFSQATHQLG-----	EGESYSM-----PTAFGL-----APTSHLE	241
QY	272	G-----ESPSPSPSPSRSSSG--	EGTCACVCGDNAACOYGVRTCEGCKGFKKTVOKNK	325
DB	242	GSGLDTP-VTSTKARGAPGGSEGRCAVCGDNASCQHYGVRTCEGCKGFKKTVOKNK	300	
QY	326	YVCLANKNCVDKRRNRNCOYCFQKCLSVGMVKEVVRTDSLKRRRRLSKPKSPLOQE	385	
DB	301	YICLANKCVDKRRNRNCOFCRFQKCLAVGMVKEVVRTDSLKRRRLSKPKQP----	356	
QY	386	PSQSPSPSPICMNALVRALTDSPT--	RDLDSRYCPTDOAAAAG-TDAASHVOOFYNLLT	442
DB	357	-----PDAFPANLLTSIVRAHLSGSPATAKLVSKFQELVPIFGKEDAGDVQOQFYDLLS	411	
QY	443	ASIDVRSWAEEKIPGFTDLPKEQDTLLIESAFLEFVLRLSIRSNTAEDKFVFCNGLVHLH	502	
DB	412	GSLEVIRKWAEEKIPGAELSPADQDLLLLESAFLFILRAYLSKPGEKGLIFCSGLVLH	471	
QY	503	RLOCILGFGFWLSDISKDFSLNLOSINLDIOALACLISALSMITERHGHLKEPRVBEELCNKI	562	
DB	472	RLOCARGFGDWIIDSIILAFSRLSHSLVVDVPAFACLSALVITDRHGLQEPFRVBEELQNRI	531	
QY	563	TSSLKHQHSKGQALEPTE-----	SKVLGALVELRKICTLGLQRIFYLKLEDLVGPSPIIDK	618
DB	532	ASCLAKEHA-AVAGEQPASCLRLLGKLELRTLCIQGLQRIFYLKLEDLVPPPPIIDK	590	
QY	619	LFLDTLPLF	626	
DB	591	IFMDTLPLF	598	